EXHIBIT 3

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 December 2001 (06.12.2001)

PCT

(10) International Publication Number
WO 01/92284 A1

(51) International Patent Classification[7]: C07H 19/10, 19/20, 21/00, C12Q 1/68

(21) International Application Number: PCT/GB01/02402

(22) International Filing Date: 31 May 2001 (31.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0013276.1 1 June 2000 (01.06.2000) GB

(71) Applicant *(for all designated States except US)*: **AMERSHAM PHARMACIA BIOTECH UK LIMITED** [GB/GB]; Amersham Place, Little Chalfont, Buckinghamshire HP7 9NA (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **ODEDRA, Raj** [GB/GB]; Amersham Pharmacia Biotech UK Ltd, Amersham Laboratories, White Lion Road, Amersham, Buckinghamshire HP7 9LL (GB). **SIMMONDS, Adrian** [GB/GB]; Amersham Pharmacia Biotech UK Ltd, Amersham Laboratories, White Lion Road, Amersham, Buckinghamshire HP7 9LL (GB). **PICKERING, Lee** [GB/GB]; Amersham Pharmacia Biotech UK Ltd, Amersham Laboratories, White Lion Road, Amersham, Buckinghamshire HP7 9LL (GB).

(74) Agents: **HAMMER, Catriona, MacLeod** et al.; Nycomed Amersham plc, Amersham Laboratories, White Lion Road, Amersham, Buckinghamshire HP7 9LL (GB).

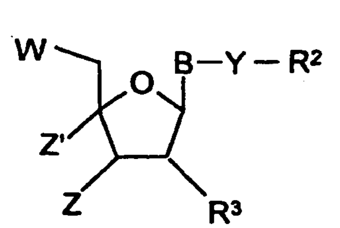
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: NUCLEOTIDE ANALOGUES COMPRISING A REPORTER MOIETY AND A POLYMERASE ENZYME BLOCKING MOIETY



(I)

(57) Abstract: Nucleotides comprising a reporter moiety and a polymerase enzyme blocking moiety in which the reporter moiety does not also act as a polymerase enzyme blocking moiety are described. Also described are compounds of Formula (I): wherein W is a phosphate group, B is a base, Y is a linker comprising an enzyme-cleavable group, R^2 is a reporter moiety, R^3 is selected from H or OH, Z and Z' are selected from H, OH, or a group X-R^1, wherein X is a linker comprising an enzyme-cleavable group and R^1 is a polymerase enzyme blocking group, provided that at least one of Z and Z' is X- R^1.

WO 01/92284 A1

Applicants: Jingyue Ju et al.
Serial No.: 10/735,081
Filed: December 11, 2003
Exhibit 3

NUCLEOTIDE ANALOGUES COMPRISING A REPORTER MOIETY AND A POLYMERASE ENZYME BLOCKING MOIETY

FIELD OF THE INVENTION

The present invention relates to nucleoside and nucleotide analogues. In particular, the invention relates to nucleotide analogues having enzyme-cleavable blocking and reporter groups positioned on separate parts of the nucleotide.

BACKGROUND OF THE INVENTION

Recent improvements in DNA sequencing techniques have sought to meet the increasing demands of large scale sequencing. Increasingly, methods in which the template nucleic acid molecules are attached to a solid surface are being developed (see, for example, US 5,302,509 and US 5,547,839). Such methods dispense with the need for an electrophoretic separation step and, with the use of optical detection technologies (see, for example, Nie et al. Annu. Rev. Biophys. Biomol. Struct. 1997, 26: 567-96), aim to allow sequencing information at the level of a single molecule to be obtained. This has the further potential for multiple samples to be analysed simultaneously.

One example of such methods is Base Addition Sequencing Scheme (BASS) (see, for example, Metzker et al., Nucleic Acids Res 1994, Vol.22, No.20; p. 4259-4267). BASS is a method involving the incorporation of nucleotide analogues which have been modified so as to comprise a blocking group which terminates DNA synthesis. A primer is annealed to a template bound to a solid support and sequence data obtained by repetitive cycles of incorporation of modified nucleotides. At each cycle, the incorporated base is identified *in situ* before being deprotected to remove the blocking group and allow the next cycle of DNA synthesis.

Methods such as BASS rely on the use of nucleotide analogues that possess polymerase enzyme blocking (or terminator) groups at the 3' hydroxyl position of the sugar on the nucleotide. Typically, the blocking group is a combined terminator and label/reporter moiety such that the incorporated nucleotide can be detected while the bulky label or reporter moiety itself fulfils the role of blocking a polymerase from any further DNA synthesis. Conveniently, as the terminator group is also the reporter

moiety, a single reaction allows simultaneous removal of both functions thus allowing subsequent DNA synthesis and for incorporation of the next base to be read.

In order to allow subsequent rounds of DNA synthesis, these polymerase enzyme blocking groups are, typically, attached to the nucleotide via a linking group in such a way that they can be removed. However, conventional sequencing strategies require high temperatures of cycling (typically approximately 95°C or above) which are associated with pH changes in the reaction mixture. Such conditions can cause reactivity of certain chemical bonds. Accordingly, the coupling methods for attaching blocking and labelling groups to nucleotides which have been used to date have focused on using those linking groups which can withstand changes in chemical conditions (such as temperature and pH). For example, the blocking and label groups can be attached via photosensitive linkage groups and thus cleavable by light irradiation (i.e. photochemical means, see, for example, WO 93/05183) or via chemical means.

However, the use of known nucleotide analogues suffers from a number of disadvantages.

Firstly, by attaching the bulky reporter moiety in the 3' position of the nucleotide, the ability of the DNA polymerase to recognise or tolerate the nucleotide is reduced. Currently known nucleotide terminators are incorporated by polymerases with an efficiency which fails to approach 97%. In addition to being poorly incorporated, modified nucleotides may be inactive (i.e. not incorporated), inhibitory (i.e. inhibit DNA synthesis) or may result in an alteration of the polymerase enzyme fidelity.

Secondly, the known methods of removing the terminator groups require repeated insult by reactive chemicals or irradiation which can result in damage to the template DNA strand through reactions such as base transformation, crosslinking, or depurination.

Any one of, or a combination of, these effects will result in a reduced accuracy in the sequence data obtained and, in particular, a decreased signal-to-noise ratio will be found on detection. Moreover, this means that the amount of sequence data that can

be obtained from successive rounds of enzyme incorporation and cleavage is limited. For example, if a combined error of approximately 3% in incorporation and cleavage were to accumulate, the result would be that sequence could only be obtained from 5 bases or fewer of the template DNA before the decreased signal to noise ratio made further sequencing impractical.

Accordingly, there is a need for improved nucleotide analogues. Such analogues may have one or more of the following attributes: tolerated by polymerases; stable during the polymerization phase; and blocking groups can be removed efficiently under conditions which minimise damage to the template strand or template-primer complex. Preferably, the improved analogues display more than one of these features and most preferably they display all of these features.

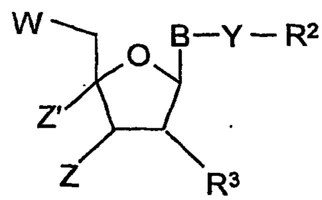
It is thus an object of the invention to provide a nucleotide analogue to which blocking and reporter moieties are attached at separate positions of the nucleotide. It is another object of the invention to provide a nucleotide analogue to which blocking and reporter moieties are attached via linking groups which are enzyme-cleavable groups. Such latter nucleotide analogues are most suitable for using in sequencing reactions which involve an isothermic reaction and therefore do not involve exposure of the nucleotide analogues to high temperatures and to undesirable variations in chemical conditions. Under the conditions of suitable sequencing reactions, including array-based sequencing technologies (such as BASS), enzyme-cleavable groups will be essentially stable. The use of enzyme-cleavable linking groups removes the need for harsh, template-damaging treatments to remove the blocking and reporter moieties.

DESCRIPTION OF THE INVENTION

The present invention describes the separation of blocking and reporter moieties on a nucleotide and the use of linkage groups cleavable by enzymatic action to attach blocking and reporter moieties to nucleotides.

Accordingly, in a first aspect, the invention provides a nucleotide comprising a reporter moiety and a polymerase enzyme blocking moiety characterised in that the reporter moiety does not also act as a polymerase enzyme blocking moiety.

In a second aspect, the invention provides a compound of Formula I:



(I)

wherein W is a phosphate group

B is a base

Y is a linker comprising an enzyme-cleavable group

R^2 is a reporter moiety

R^3 is selected from H or OH

Z and Z' are selected from H, OH, or a group X-R^1, wherein X is a linker comprising an enzyme-cleavable group and R^1 is a polymerase enzyme blocking group, provided that at least one of Z and Z' is X- R^1.

Suitably, W represents a phosphate group and may be a mono-, di- or tri- phosphate group. In a particularly preferred embodiment, W is a triphosphate.

Suitable bases, B, include purines or pyrimidines and, in particular, any of the bases A, C, G, U and T or their analogues.

Suitably, only one of Z and Z' is X-R^1. In one preferred embodiment, Z is X-R^1 and Z' is H or OH.

In a preferred embodiment, X and/or Y may be a chain of up to 30 bond lengths and may include atoms selected from carbon, nitrogen, oxygen and sulphur atoms, the linker group may be rigid or flexible, unsaturated or saturated as is well known in the field. X and/or Y may further incorporate one or more amino acids joined by peptide

bonds. The incorporation of amino acids can be through the incorporation of amino acid monomers or oligomers using standard amino acid chemistry (see, for example, "Synthetic Peptides - A Users Guide" Ed. G. A. Grant; 1992). Suitably, linker Y links the base, B to the reporter moiety R^2.

Suitable enzyme-cleavable groups in X and Y include any chemical structure which is recognisable by an enzyme and which, in an enzyme-cleavage reaction, results in the polymerase enzyme blocking group (R^1) and/or the reporter moiety (R^2) being detached from the compound. The enzyme-cleavable groups in X and Y can be the same or different. In one embodiment, X and/or Y incorporate amino linkage groups.

In a preferred embodiment, where Z is X-R^1, the enzyme-cleavage reaction results in the formation of an –OH group in the 3' position thus leaving the incorporated nucleotide capable of binding to a subsequent nucleotide.

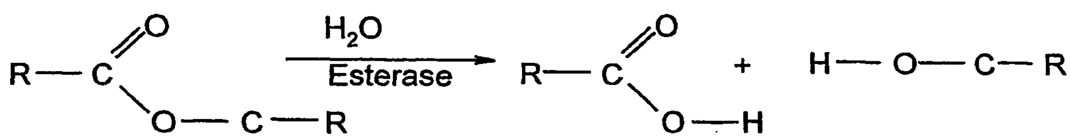
In another embodiment, where Z' is X-R^1, the enzyme-cleavage reaction leaves a group in the 4' position which allows subsequent incorporation of a nucleotide (i.e. chain extension). In a particularly preferred embodiment, the enzyme cleavage reaction leaves an amino methyl group in the 4' position.

In a particularly preferred embodiment, X and Y comprise the same enzyme-cleavable group thus facilitating a single addition or reaction causing cleavage of both blocking and reporter groups in one reaction.

In a preferred embodiment, the enzyme-cleavable groups may be, for example, groups cleavable by enzymes such as esterases, phosphatases, peptidases (i.e. endo or exo peptidases), amidases, glucosidases or phosphorylases. Suitable enzymes are those that are reactive under mild conditions. (see Handbook of Proteolytic Enzymes, Barrett et al., ISBN 0-12-079370-9). In a particularly preferred embodiment, the enzyme-cleavable group is cleavable by penicillin amidase.

Esterases catalyse the general reaction set out below in Reaction Scheme 1:

Reaction Scheme 1



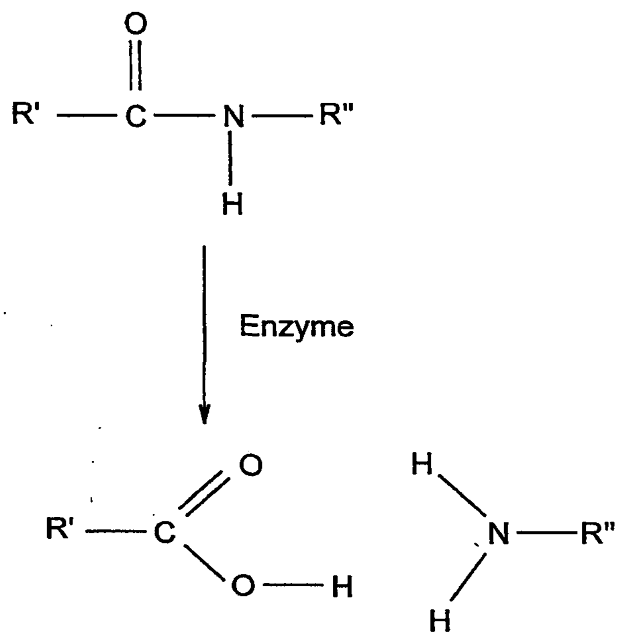
Thus, in a further preferred embodiment of the second aspect, at least one of X and Y comprise a carboxyl group.

Non-specific esterase activity is associated with a number of enzyme systems. This activity has been associated with both physiological function and drug metabolism. Such a non-specific carboxylesterase activity can be used to modify molecules in vitro. Thus in a preferred embodiment, once the nucleotides are incorporated, the linkage groups may be digested with a non-specific esterase to remove the blocking group and reporter moiety without damaging the template strand or the template/primer complex. Following deprotection, DNA synthesis is reinitiated leading to the next cycle of labelled analogue addition.

Other suitable enzyme-cleavable groups include those cleavable by amidases and peptidases.

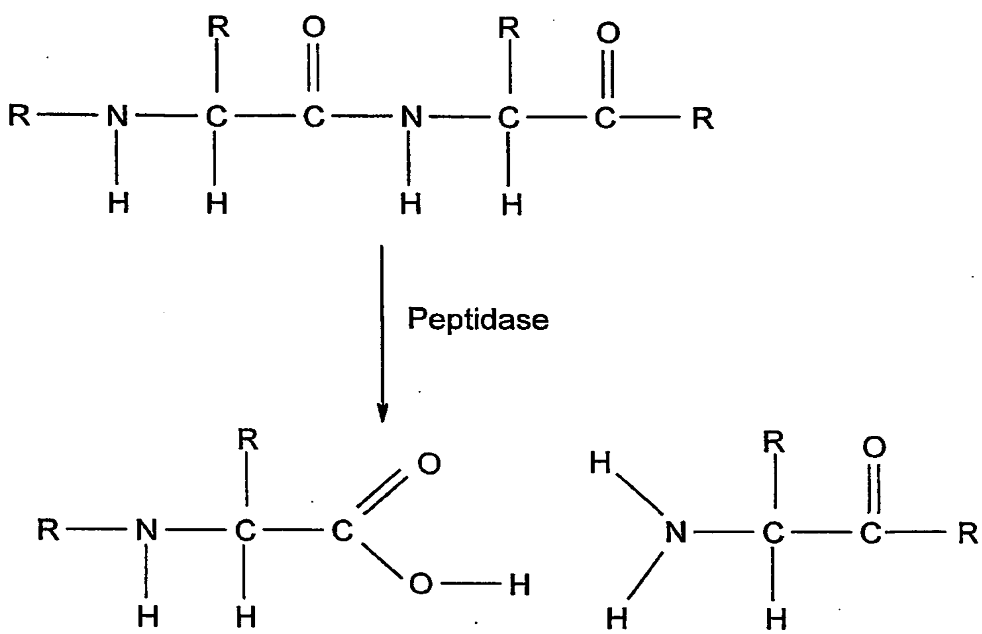
Amidases catalyse the cleavage of amide bonds as set out in Reaction Scheme 2.

Reaction Scheme 2

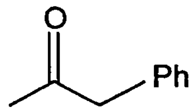


Where R' and R" both represent one or more amino acid residues, then peptidases catalyse the following general reaction set out in Reaction Scheme 3:

Reaction Scheme 3



Penicillin amidase (also known as penicillin aminohydrolase; EC 3.5.1.11) cleaves the group:



In addition, cleavage of phenylacetic acid by penicillin amidase has been described in WO 97/20855. Accordingly, in a particularly preferred embodiment, X and/or Y comprise a penicillin amidase cleavage site.

Suitable methods for attaching a linker comprising an enzyme cleavable group to a base moiety are described, for example, in Cavallaro et al. Bioconjugate Chem. 2001, 12, 143-151. Further methods are described in Langer et al, Proc Natl Acad Sci USA, 1981, 78, 6633-6637; Livak et al, Nucleic Acids Res, 1992, 20, 4831-4837 and Gebeyehu et al, Nucleic Acids Res, 1987, 15, 4513-4534.

A suitable reporter moiety, R^2, may be any one of various known reporting systems. It may be a radioisotope by means of which the nucleoside analogue is rendered easily detectable, for example ^{32}P, ^{33}P, ^{35}S incorporated in a phosphate or thiophosphate or H phosphonate group or alternatively ^{3}H or ^{14}C or an iodine isotope. It may be an isotope detectable by mass spectrometry or NMR. It may be a signal moiety e.g. an enzyme, hapten, fluorophore, chromophore, chemiluminescent group, Raman label, electrochemical label, or signal compound adapted for detection by mass spectrometry.

In a preferred embodiment, the reporter moiety has fluorescent properties and can be detected using a sensitive fluorescence detector. It may be a fluorophore, for example, selected from fluoresceins, rhodamines, coumarins, BODIPY™ dyes, cyanine dyes and squarate dyes (described, for example, in WO 97/40104). Most preferably, the reporter moiety is a cyanine dye. The Cyanine dyes (sometimes referred to as "Cy dyes™"), described, for example, in US Patent 5,268,486, is a series of biologically compatible fluorophores which are characterised by high fluorescence emission,

environmental stability and a range of emission wavelengths extending into the near infra-red which can be selected by varying the internal molecular skeleton of the fluorophore.

The reporter moiety may comprise a signal moiety and a linker group joining it to the remainder of the molecule, which linker group may be a chain of up to 30 bond lengths and may include atoms selected from carbon, nitrogen, oxygen and sulphur atoms, the linker group may be rigid or flexible, unsaturated or saturated as is well known in the field.

In a preferred embodiment, different reporter moieties will be chosen such that more than one base can be incorporated and detected in a single sequencing reaction. In a particularly preferred embodiment, each base will be labelled with a different reporter moiety so that all four bases can be used at the same time in a sequencing reaction. The different reporter moieties will enable the different bases to be distinguishable by fluorescence spectroscopy or other optical means. In a preferred embodiment, the reporter moiety is chosen such that 4 distinguishable moieties can be used to label each of the 4 natural bases, A, G, C and T or their analogues such that each of the nucleotides are distinguishable from each other.

Suitably R^3 is selected from H or OH. Thus ribonucleotides and deoxyribonucleotides are envisaged together with other nucleoside analogues.

A polymerase enzyme blocking group, R^1, is one which should have the functional properties of blocking further elongation of the polymer once the nucleotide of the present invention has been incorporated by a selected polymerase in selected polymerase enzyme conditions. In particular, a blocking group is any chemical group which can be attached to a nucleotide and which will allow the 5' end of the modified nucleotide to attach to a 3' end of another nucleotide in a DNA chain but will not allow attachment of a nucleotide to the 3'hydroxyl group of the modified nucleotide. Suitably, the absence of an OH group in the 3' position will prevent further elongation by polymerase activity. In a particularly preferred embodiment, the blocking group, R^1 is selected from acetyl, CH_3, glycyl, leucyl and alanyl groups. In another embodiment, the blocking group may be in the form of a di or tri peptide.

In another embodiment, the polymerase enzyme blocking group can be attached at the 4' position i.e. Z' is X-R^1. It is postulated that modification at this position results in an analogue that is more readily accepted as a substrate by polymerases. Methods for synthesising nucleotide analogues having a 4' blocking modification are described by Giese et al. in EP 0,799,834.

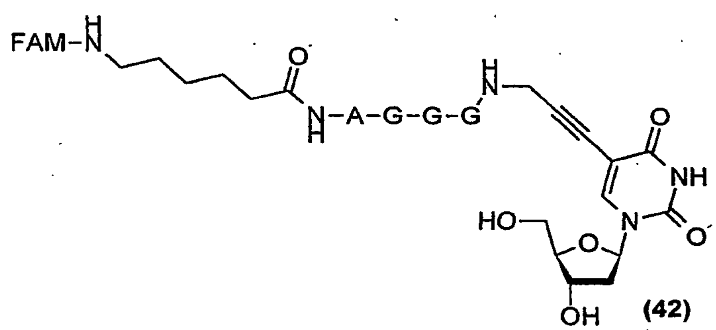
Where Z' is X-R^1, Z is, preferably, OH. In this preferred embodiment, it is speculated that the presence of a 3' OH group in the modified nucleotide of Formula I may facilitate polymerase recognition of the modified analogue and would also leave the nucleotide ready for subsequent incorporation upon detaching blocking group, R^1.

Preferably, the blocking group R^1 does not also act as a reporter molecule.

In another preferred embodiment, the modified nucleotide remains compatible with elongation enzymology, i.e. it can still be incorporated by a polymerase. Procedures for selecting suitable nucleotide and polymerase combinations will be readily adapted from Metzker et al, Nucleic Acids Res 1994, Vol. 22, No. 20, 4259-4267. In particular, it is desired that a selected polymerase be capable of selectively incorporating a nucleotide.

Examples of particularly preferred compounds of Formula I, in which R^2 is the cyanine dye Cy3, include 5-{[N-(carboxypentyl)-trimethinecyanine]amidoacetic acid-3-allyl ester}-3'-acetoxy-2'-deoxyuridine -5'-triphosphate, 5-{[N-(carboxypentyl)-trimethinecyanine]amidoacetic acid-3-allyl ester}-3'-acetoxy-2'-deoxycytidine -5'-triphosphate, 7-{[N-(carboxypentyl)-trimethinecyanine]amidoacetic acid-3-allyl ester}-3'-acetoxy-2'-deoxy-7-deazaadenosine -5'-triphosphate and 7-{[N-(carboxypentyl)-trimethinecyanine]amidoacetic acid-3-allyl ester}-3'-acetoxy-2'-deoxy-7-deazaguanosine -5'-triphosphate. These preferred compounds are shown in Figure 2. A further preferred compound is 5-[*N*-Fluorescein-5(and-6) carboxamidohexanoyl]propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine triphosphate (shown in Figure 6).

In a third aspect of the invention, there is provided a chemical intermediate selected from the group consisting of: 4'-*C*-(Glycylaminomethyl)thymidine triphosphate; 4'-*C*-(*N*-trifluoroacetylaminomethyl)thymidine triphosphate; 4'-*C*-(Aminomethyl)thymidine triphosphate (shown in Figure 3); 5-(*N*-Trifluoroacetyl)propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine; 4'-*C*-(N-Acetylglycylaminomethyl)thymidine; 4'-*C*-(*N*-Leucylaminomethyl)thymidine (shown in Figure 7); 4'-*C*-(*N*-Glycylaminomethyl)thymidine (shown in Figure 8); *N*-{α-[4'-methyloxythymidyl]phenyl}phenylacetamide triphosphate (shown in Figure 9); *N*-[α-(3'-*O*-thymidyl)-phenyl)]-phenylacetamide triphosphate, *N*-{α-[3'-*O*-(5- *N*-(α-methyloxy-*N*'-trifluoroacetylaminopropyl benzamide) phenylacetamide –2'-deoxyuridyl]-phenyl}-phenylacetamide (shown in Figure 11) and a compound of the following formula:



(42)

In another aspect of the invention, there is provided a process for the manufacture of a compound in accordance with any of the first, second or third aspects. In a further aspect there is provided a process for the manufacture of a compound of Formula I using an intermediate compound as defined in the third aspect.

In a fourth aspect of the invention, there is provided, a set of nucleotides characterised in that the set contains at least one compound of Formula I. Preferably, such a set will comprise each of the four natural bases A, G, C and T (or their analogues) wherein at least one is a compound of Formula I.

In a preferred embodiment of the fourth aspect the set of nucleotides will comprise at least two compounds of Formula I having different bases, B, characterised in that each

compound of Formula I has a different reporter moiety, R^2. Thus, for example, the set of nucleotides may comprise compounds of Formula I with bases A and G wherein the compound with base, A, has a first reporter moiety (R^2) and the compound with base, G, has a second reporter moiety (R^2) wherein the first and second reporter molecules are distinguishable from each other.

In another preferred embodiment of the fourth aspect, the set of nucleotides comprises four compounds of Formula I characterised in that each compound has a different base, B, such that each of the bases A, G, C and T are present and each of the four compounds of Formula I has a reporter moiety which is distinguishable from all of the other three bases.

In a fifth aspect of the invention there is provided, a method for nucleic acid molecule sequencing comprising the steps of

a) immobilising a complex of a primer and a template to a solid phase
b) incubating with a polymerase in the presence of a compound of Formula I.

In one embodiment of the fifth aspect, the complex of primer and template can be preformed by incubation under appropriate hybridisation conditions before immobilising the complex onto a solid phase. In another embodiment, the primer or the template can be immobilised onto a solid phase prior to formation of the complex by introduction of the appropriate hybridisation partner (i.e. template or primer, respectively). In yet another embodiment, the complex immobilised onto the solid phase can be a single nucleic acid molecule comprising both "primer" and "template"; for example, the immobilised nucleotide can be a hairpin structure.

Suitable polymerases are enzymes that perform template-dependent base addition including DNA polymerases, reverse transcriptases and RNA polymerases. Suitable native or engineered polymerases include but are not limited to T7 polymerase, the Klenow fragment of E. coli polymerase which lacks 3'-5'exonuclease activity, E. coli polymerase III, Sequenase™, ϕ29 DNA polymerase, exonuclease-free Pfu, exonuclease-free Vent™ polymerase, Thermosequenase, Thermosequenase II, Tth DNA polymerase, Tts DNA polymerase, MuLv Reverse transcriptase or HIV reverse

transcriptase. The selection of an appropriate polymerase depends on the interaction between a polymerase and the specific modified nucleotide (as described by Metzker et al., Nucleic Acids Res 1994, Vol.22, No.20; p. 4259-4267).

Nucleotides comprising enzyme-cleavable linkage groups such as carboxyl ester attachment groups are suitable for use in sequencing reactions used in array based sequencing, such as BASS. Such reactions are isothermic, unlike cycle sequencing, so allowing much better control of reaction conditions. In particular, the sequencing reaction takes place at relatively low temperatures (typically less than 70°C) thus enabling enzyme-cleavable linkage groups, such as the carboxyl ester attachment, to remain stable under these sequencing reaction conditions. Accordingly, polymerases which may be useful in the fifth aspect of the invention include thermostable polymerases and non-thermostable polymerases.

In a preferred embodiment of the fifth aspect, the method further comprises the steps of

c) detecting the incorporation of a compound of Formula I
d) incubating in the presence of enzyme under suitable conditions for enzymatic cleavage of the enzyme-cleavable groups X and Y

Suitable conditions for enzyme cleavage of the enzyme-cleavable groups will depend on the nature of the enzymes involved. Enzymes such as carboxyesterases are active under a broad range of conditions and do not require co-factors. Commercially available carboxyesterases will hydrolyse esters under mild pH conditions of between pH 7.0 and pH 8.0. e.g. 0.1M NaCl, 0.05M Tris.HCl, pH 7.5. Suitable conditions for cleavage by amidases and peptidases are exemplified in Example 8 below.

In another embodiment of the fifth aspect, the method further comprises
e) repeating steps a)-d)

In a preferred embodiment of the fifth aspect, the enzyme in step d) is an amidase.

In a further aspect of the invention there is provided use of a compound of Formula I in a sequencing reaction.

Briefly, sequencing reactions using modified nucleotides in accordance with the first aspect of the invention may be performed as follows. Primer template complexes are immobilised to a solid surface and contacted with modified nucleotides in the presence of a suitable buffer also containing a polymerase, such as Klenow fragment of E. coli polymerase which lacks 3'-5'exonuclease activity, and a commercially available pyrophosphatase. The reaction is incubated under suitable conditions for a polymerase-mediated base addition reaction followed by the removal of non-incorporated nucleotides and enzymes by washing with a wash buffer. Suitably, the wash buffer contains a buffering agent, such as an organic salt, to maintain a stable pH of approximately pH 6 to pH 9 and possibly also comprises monovalent or divalent cations and a detergent so as to eliminate non-covalently bound molecules from the solid surface. Where the modified nucleotides comprise a fluorescent reporter molecule, incorporated nucleotides are detected by measuring fluorescence and the corresponding nucleotide identified. Following identification, the templates are contacted with a buffered solution containing an excess of a protein displaying the appropriate enzyme activity and incubated under conditions for enzyme cleavage activity. For example, where the enzyme-cleavable group linking reporter molecule and/or blocking group to the nucleotides is a carboxyl group, the solution contains an excess of a protein displaying non-specific esterase activity. Following enzyme activity, the products of enzymatic cleavage are eliminated by washing as above. Following the washing step, the immobilised template is washed with an excess of buffer used for the polymerase reaction and the steps of polymerase-mediated base addition, detection of incorporated nucleotide and enzyme-cleavage activity are repeated to obtain further sequence data.

SPECIFIC DESCRIPTION

For the purposes of clarity, certain embodiments of the present invention will now be described by way of example with reference to the following figures:

Figure 1 shows a reaction scheme for synthesising a compound of Formula I.
Figure 2 shows examples of compounds of Formula I.
Figure 3 shows a reaction scheme for synthesising nucleotide analogues.

Figures 4 and 5 show reaction schemes for synthesising nucleoside analogues.
Figure 6 shows a reaction scheme for synthesising a fluorescein-labelled nucleotide.
Figures 7 and 8 show reaction schemes for synthesising nucleoside analogues.
Figures 9 and 10 show reaction schemes for synthesising nucleotide analogues.
Figure 11 shows a reaction scheme for synthesising a nucleoside analogue.
Figure 12 shows a reaction scheme for synthesising a fluorescein-labelled nucleoside.

Example 1

A reaction scheme for the synthesis of an example of a compound of Formula I is set out in Figure 1 using 5-iodo-2'-deoxyuridine (Sigma-Aldrich Chemical Co.) as the starting material and incorporating the dye, Cy3, by reaction with (N-hydroxy)succinimide Cy3 ester.

In the reaction scheme, the following abbreviations are used: TBDMS – tertButyldimethylsilyl; DMTr – Dimethoxytrityl; TFA – trifluoroacyl; nBu - linear Butyl chain; MeCN – Acetonitrile; TCA - Trichloroacetic acid; DCM – Dichloromethane; DMF - N,N-Dimethylformamide; TEAB - Tetraethylammonium bicarbonate buffer; THF – Tetrahydrofuran; Ac – Acetyl.

Wherein TBDMS, Ac, TFA and nBu have the following structural formulae:

TBDMS Ac TFA nBu

Example 2

Synthesis of nucleotides with a 4' groups (4'-*C* (Glycylaminomethyl) thymidine triphosphate **(6)**, 4'-*C*-(*N*-trifluoroacetylaminomethyl)thymidine triphosphate **(8)** and 4'-*C*-(Aminomethyl)thymidine triphosphate **(9)**).

Figure 3 illustrates the synthetic pathway for nucleotide analogues indicated as **(6)**, **(8)** and **(9)** via the intermediate compound **(4)**.

i) 3'-*O*-[(*tert*butyl)dimethylsilyl]-5'-*O*-[(*tert*butyl)diphenylsilyl]-4'-*C*-(1,1'-diphenylmethylaminomethyl)thymidine (2).

The starting material 3'-*O*-[(*tert*butyl)dimethylsilyl]-5'-*O*-[(*tert*butyl)diphenylsilyl]-4'-*C*-formylthymidine **(1)** (described in Marx et al. *Helv. Chim. Acta.* 1996, **79**, *1980 – 1994* and references cited therein.) (1.7g, 2.7mmol) and Ph_2CHNH_2 (0.53g, 2.9mmol, 0.5mL) were dissolved in anhydrous acetonitrile (10mL) at ambient temperature. Neat glacial acetic acid (0.11g, 1.9mmol, 0.11mL) was then added and the solution stirred for 30 minutes at ambient temperature, during which time a white precipitate formed. Solid $NaBH_3CN$ (0.25g, 4.1mmol) was then added in portions to the stirred suspension. The resulting mixture was then stirred at ambient temperature for 18 hours. The solvent was then removed under vacuum and the residue redissolved in dichloromethane and washed with brine. The organic layer was then separated, dried over $MgSO_4$ and then filtered. Concentration of the filtrate under vacuum gave a white foam which was purified by flash column chromatography (7:3, dichloromethane:ethyl acetate) to afford the title compound **(2)** (1.2g, 55%) as a white foam. δ_H (300MHz, $CDCl_3$) 8.3(1H, s, br, N^3-H), 7.72 – 7.65(4H, dd, Ph), 7.49 – 7.14(17H, m, <u>Ph_2</u>CHNH, <u>Ph_2</u>Si*t*Bu; H-6, s, obs), 6.37(1H, dd, H-1'), 4.67(1H, s, Ph_2<u>CH</u>NH), 4.57(1H, dd, H-3'), 3.96(2H, 2d, H-5'), 2.70(1H, d, 4'-*C*-CH_2), 2.55(1H, d, 4'-*C*-CH_2), 2.25(2H, m, H-2'), 1.61(3H, s, 5-CH_3), 1.11(9H, s, <u>*t*Bu</u>$SiPh_2$), 0.75(9H, s, <u>*t*Bu</u>$SiMe_2$), 0.11(3H, s, MeSi), 0.01(3H, s, MeSi); δ_C (75.45MHz, $CDCl_3$) 163.47, 150.13, 144.14, 143.79, 135.72, 135.65, 133.07, 132.58, 130.14, 129.63, 128.67, 128.59, 128.41, 127.97, 126.94, 126.88, 111.03, 89.79, 83.84, 72.89, 67.99, 66.74, 49.80, 41.85, 27.08, 25.63, 19.44, 17.85, 14.20, 12.07, -4.68, -5.36; ν cm^{-1} 1688, 1112, 833.

ii) 4'-*C*-(1,1'-diphenylmethylaminomethyl)thymidine (3).

3'-*O*-[(*tert*butyl)dimethylsilyl]-5'-*O*-[(*tert*butyl)diphenylsilyl]-4'-*C*-(1,1'-diphenylmethylaminomethyl)thymidine **(2)** (1.2g, 1.5mmol) was dissolved in tetrahydrofuran (50mL) at ambient temperature. A solution of tetrabutylammonium fluoride in tetrahydrofuran (4.5mL) was then added to the solution. The resulting mixture was stirred for 3 hours at ambient temperature. Methanol (5mL) was then added and the reaction mixture concentrated under vacuum. The residue was purified by flash column chromatography (95:5 dichloromethane: methanol) to give the title

compound **(3)** as an amorphous white solid (0.65g, 100%) on removal of solvent from the appropriate fractions. δ_H (300MHz, d_6-DMSO) 11.23(1H, s, br, N^3-H), 7.75(1H, s, H-6), 7.41 – 7.15(10H, m, $\underline{Ph_2}$CHNH), 6.21(1H, dd, H-1'), 5.57(1H, d, br, 3'-OH), 5.21(1H, t, 5'-OH), 4.78(1H, s, $Ph_2\underline{CH}$NH), 4.35(1H, m, H-3'), 3.61(2H, 2d, H-5'), 2.54(2H, 2d, 4'-*C*-CH_2), 2.18(2H, m, H-2'), 1.75(3H, s, 5-CH_3); δ_C (75.45MHz, d_6-DMSO) 150.46, 144.53, 144.28, 136.28, 128.36, 126.99, 126.74, 109.17, 88.39, 83.18, 66.79, 64.10, 54.91, 12.27.

iii) 4'-*C*-Aminomethylthymidine (4).

4'-*C*-(1,1'-diphenylmethylaminomethyl)thymidine **(3)** (0.65g, 1.5mmol) was dissolved in absolute ethanol (50mL). Cyclohexene (12mL), glacial acetic acid (5mL) and 5%-palladium on charcoal (0.6g) were added and the mixture heated under reflux with stirring. The solution was heated for 5 hours and then allowed to cool to ambient temperature. Palladium catalyst was removed by filtration through celite and the filtrate was concentrated under vacuum to give the crude *title* compound **(4)**, which was used in the following step without further purification. ES +ve *m/z* 272$(M+H)^+$, 242$(M-CH_2NH_2)^+$. For an alternative synthesis of this compound see: Wang et al; *Tetrahedron Letters* 1996, **37**, *6515 – 6518*.

iv) 4'-*C*-(N-trifluoroacetylglycylaminomethyl)thymidine (5).

4'-*C*-Aminomethylthymidine **(4)** (0.14mmol) was dissolved in anhydrous *N,N*-dimethylformamide (1mL). *N*-Trifluoroacetylglycine (0.036g, 0.21mmol) and *O*-(*N*-succimidyl)-*N,N,N',N'*-tetramethyluronium tetrafluoroborate [TSTU] (0.085, 0.28mmol) were then added and the solution stirred at ambient temperature. Diisopropylethylamine (0.036g, 0.28mmol, 0.05mL) was then added dropwise to the solution with stirring. After 12 hours the solvent was removed under vacuum and the residue redissolved in dichloromethane: methanol (9:1). Applied solution directly to a silica gel column and eluted with dichloromethane: methanol (9:1). Fractions containing the desired material were pooled and the solvent removed under vacuum to give the title compound **(5)** as a clear, glassy material (0.28g, 28% from crude **4**). δ_H (300MHz, CD_3OD) 7.91(1H, s, H-6), 6.27(1H, dd, H-1'), 4.49(1H, dd, H-3'), 3.74(2H, s, glycyl CH_2), 3.52(1H, d, H-5'), 3.51(2H, s, 4'-*C*-CH_2), 3.43(1H, d, H-5'),

2.32(2H, m, H-2'), 1.68(3H, s, 5-CH_3); ES +ve *m/z* 425$(M+H)^+$, 442$(M+H_2O)^+$, 447$(M+Na)^+$.

v) 4'-*C*-(Glycylaminomethyl)thymidine triphosphate (6).

4'-*C*-(N-trifluoroacetylglycylaminomethyl)thymidine **(5)** (0.026g, 0.06mmol) and proton sponge (0.02g, 0.09mmol) were dissolved in trimethylphosphate (1mL) at ambient temperature. The solution was then cooled to 0C on an ice bath under an atmosphere of nitrogen. Phosphorous oxychloride (0.02g, 0.12mmol, 0.01mL) was then added dropwise to the cooled mixture while stirring. After stirring for two hours at 0C a solution of 0.5M tri-n-butylammonium pyrophosphate in *N,N*-dimethylformamide (1mL) and tributylamine (0.075mL) was added to the cooled solution. After stirring for a further two minutes 0.2M triethylammonium bicarbonate buffer (2mL) was added and the solution allowed to warm to ambient temperature. The solution was stirred for 45 minutes at ambient temperature and then concentrated under vacuum. The oily residue was then redissolved in concentrated ammonia solution (1mL) and allowed to stand overnight at ambient temperature. Removal of the solvent under vacuum then gave an oily residue which was redissolved in distilled water and subjected to purification by ion exchange chromatography (DEAE Sephacryl, eluted with 0 – 100% water – 0.8M triethylammonium bicarbonate buffer). Fractions containing the desired product were detected by U.V. and pooled. Lyophilisation of the appropriate fractions gave a white residue which was redissolved in water and eluted through a preparative C18 hplc column with water. Fractions containing the desired product were pooled and lyophilised to give the title compound **(6)** as a white foam (0.8μmol, 13%). δ_H (300MHz, D_2O) 7.48(1H, s, H-6), 6.19(1H, dd, H-1'), 3.99(3H, dd, H-3',m, br, H-5'), 3.50(1H, d, 4'-*C*-CH_2), 3.38(2H, s, glycyl CH_2), 3.36(1H, d, 4'-*C*-CH_2), 2.36(2H, m, H-2'), 1.77(3H, s, 5-CH_3); δ_P (121.5MHz, D_2O) –5.6(d), -10.8(d), -19.0(t); λ_{max} 264nm.

vi) 4'-*C*-(N-trifluoroacetylaminomethyl)thymidine (7).

4'-*C*-Aminomethylthymidine **(4)** (1.5mmol) was dissolved in methanol (10mL) and triethylamine (1.5mL) at ambient temperature. The solution was then treated with neat ethyl trifluoroacetate (1mL) and the resulting solution stirred for 18 hours at ambient temperature with exclusion of moisture. The solvents and reagents were then

removed under vacuum to give the crude product. Purification by flash column chromatography (9:1 dichloromethane:methanol) afforded the title compound **(7)** as a white foam (0.37g, 66%). δ_H (300MHz, CD_3OD) 8.77(1H, t, br, N^3-H), 7.73(1H, s, H-6), 6.24(1H, dd, H-1'), 4.56(1H, dd, H-3'), 3.6(3H, s, H-5', obs d, 4'-*C*-CH_2), 3.53(1H, d, 4'-C-CH_2), 2.35(2H, m, H-2'), 1.81(3H, s, 5-CH_3); δ_C (75.45MHz, CD_3OD), 166.24, 160.08 – 158.62(q, CF_3CO), 152.21, 138.11, 123.01 – 111.51(q, CF_3CO), 111.51, 88.93, 85.69, 72.79, 64.23, 41.70, 40.81, 12.40; ES +ve *m/z* 368 $(M+H)^+$, 390$(M+Na)^+$; ν cm^{-1} 1684. For an alternative synthesis of this compound see Wang et al, *Tetrahedron Letters* 1996, **37**, *6515 – 6518*.

vii) 4'-*C*-(*N*-trifluoroacetylaminomethyl)thymidine triphosphate(8).

4'-*C*-(*N*-trifluoroacetylaminomethyl)thymidine **(7)** was weighed into a round-bottom flask and the flask flushed with argon. Pyridine and 1,4-dioxane were then added to give a clear solution. A 1M solution of 2-Chloro-4H-1,3,2-dioxaphosphorin-4-one was then added to the stirred solution. After stirring for 30 minutes a 0.5M solution of tributylammonium pyrophosphate in DMF and tributylamine were added together and the resulting solution stirred for a further 30 minutes. A 1% solution of iodine in pyridine: water 98:2 was then added and the solution stirred for a further 30 minutes. Excess iodine was then destroyed by the addition of saturated sodium thiosulfate solution. The reaction mixture was then concentrated to give a crude product which was then initially purified by the procedure described for the preparation of compound **(6)**. Final purification was achieved using preparative reverse-phase HPLC, eluting with 0.1M triethyammonium bicarbonate : acetonitrile. Obtained the title compound **(8)** as a white powder after lyophilisation of the appropriate fractions (0.54mg). δ_H (300MHz, D2O) 7.45(1H, s, H-6), 6.23(1H, dd, H-1'), 3.76(1H, dd, H-3'), 3.00(2H, 2d, H-5'), 2.61(2H, 2d, 4'-C-CH2), 2.10(2H, m, H-2'), 1.62(3H, s, 5-CH3); δ_P (121.5MHz, D2O) –4.9(d), -10.5(d), -20.4(t); λ_{max} 266nm.

viii) 4'-*C*-(Aminomethyl)thymidine triphosphate (9).

4'-*C*-(*N*-trifluoroacetylaminomethyl)thymidine triphosphate **(8)** was dissolved in concentrated aqueous ammonia solution and allowed to stand for 18 hours at ambient temperature. The solution was then lyophilised to give the title compound **(9)** as a white powder.

Example 3

Synthesis of nucleotide with a 4' blocking group and a fluorescent label attached to the base (5-[*N*-Fluorescein-5(and-6)carboxamidohexanoyl] propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine triphosphate **(18)**).

Figure 4 illustrates the synthetic pathway for the nucleoside analogue, 5-(*N*-Trifluoroacetyl)propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine **(16)**.

i) 3'-*O*-[(*tert*butyl)dimethylsilyl]-5'-*O*-[(*tert*butyl)diphenylsilyl]-4'-*C*-(1,1'-diphenylmethylaminomethyl)-2'-deoxyuridine (11).

3'-*O*-[(*tert*butyl)dimethylsilyl]-5'-*O*-[(*tert*butyl)diphenylsilyl]-4'-*C*-formyl-2-deoxyuridine **(10)** (Yang et al; *Tetrahedron Letters* 1992, **33**, *37 – 40.*) (0.15g, 0.25mmol) was treated with Ph_2CHNH_2 (0.07g, 0.37mmol, 0.064mL), glacial acetic acid (0.015g, 0.26mmol, 0.015mL) and $NaBH_3CN$ (0.023g, 0.37mmol) in acetonitrile (1mL) according to the procedure used for the preparation of compound **(2)**. The title compound **(11)** was obtained as clear, viscous oil (0.13g, 69%). δ_H (300MHz, $CDCl_3$) 8.02(1H, s, br, N^3-H), 7.81(1H, d, H-6), 7.69 – 7.18(20H, m, $\underline{Ph_2}$CHNH, *t*BuSi$\underline{Ph_2}$), 6.32(1H, dd, H-1'), 5.30(1H, d, H-5), 4.67(1H, s, $Ph_2\underline{CH}$NH), 4.61(1H, dd, H-3'), 4.02(1H, d, H-5'), 3.87(1H, d, H-5'), 2.69(1H, d, 4'-*C*-CH_2), 2.53(1H, d, 4'-*C*-CH_2), 2.37(1H, m, H-2'), 2.17(1H, m, H-2'), 1.10(9H, s, <u>*t*Bu</u>SiPh_2), 0.75(9H, s, <u>*t*Bu</u>SiMe_2), 0.01(6H, s, *t*BuSi$\underline{Me_2}$); ES +ve *m/z* 776$(M+H)^+$.

ii) 4'-*C*-(1,1'-diphenylmethylaminomethyl)-2'-deoxyuridine (12).

3'-*O*-[(*tert*butyl)dimethylsilyl]-5'-*O*-[(*tert*butyl)diphenylsilyl]-4'-*C*-(1,1'-diphenylmethylaminomethyl)-2'-deoxyuridine **(11)** (0.13g, 0.16mmol) was treated with tetrabutylammonium fluoride solution (0.45mL) in tetrahydrofuran (5mL) according to the procedure used for the preparation of compound **(3)**. The title compound **(12)** was obtained as a clear glass (0.056g, 78%). δ_H (300MHz, d_6-DMSO) 11.25(1H, s, br, N^3-H), 7.90(1H, d, H-6), 7.40 – 7.15(10H, m, $\underline{Ph_2}$CHNH), 6.19(1H, dd, H-1'), 5.93(1H, d, br, 3'-OH), 5.91(1H, s, br , Ph_2CH<u>NH</u>), 5.11(1H, t, br, 5'-OH), 4.78(1H, s, $Ph_2\underline{CH}$NH), 4.33(1H, dd, H-3'), 3.61(1H, d, H-5'), 2.52 – 2.49(2H,

m, 4'-C-CH_2), 2.18(2H, m, H-2'); δ_C (75.45MHz, d_6-DMSO) 163.13, 150.43, 144.50, 144.25, 140.62, 128.34, 126.97, 126.92, 126.72, 101.60, 88.63, 83.50, 72.0, 66.85, 58.8, 51.15; ES +ve *m/z* 424$(M+H)^+$.

iii) 4'-*C*-Aminomethyl-2'-deoxyuridine (13).

4'-*C*-(1,1'-diphenylmethylaminomethyl)-2'-deoxyuridine **(12)** (0.056g, 0.13mmol) was treated with cyclohexene (1mL), glacial acetic acid 0.5mL) and 5% palladium on charcoal (0.06g) in ethanol (5mL) according to the procedure used for the preparation of compound **(4)**. The crude product was purified by dissolving in water and washing with toluene. The aqueous layer was separated and the solvent removed under vacuum to give the title compound **(13)** as an amorphous solid (0.034g, 100%). δ_H (300MHz, D_2O) 7.64(1H, d, H-6), 6.22(1H, dd, H-1'), 5.72(1H, d, H-5), 4.43(1H, m, H-3'), 3.56(2H, 2d, H-5'), 3.26(1H, d, $\underline{CH_2}NH_2$), 3.05(1H, d, $\underline{CH_2}NH_2$), 2.40(2H, m, H-2'); δ_C (75.45MHz, D_2O) 166.95, 152.46, 142.67, 103.11, 86.77, 86.43, 73.97, 64.82, 41.59, 39.25; ES +ve *m/z* 258$(M+H)^+$, 242$(M-NH_3)^+$.

iv) 4'-*C*-(Acetylaminomethyl)-2'-deoxyuridine (14)

4'-*C*-Aminomethyl-2'-deoxyuridine **(13)** (0.034g, 0.13mmol) was dissolved in anhydrous *N,N*-dimethylformamide (0.5mL) at ambient temperature. Excess pentafluorophenyl acetate (0.88g, 0.39mmol) was then added directly to the reaction mixture as a solid. The resulting solution was stirred for 18 hours at ambient temperature. The solvent was then removed under vacuum and the residue re-dissolved in water and then washed with toluene. The aqueous layer was then separated and the solvent removed under vacuum. The title compound was obtained as an amorphous solid (0.032g, 82%). δ_H (300MHz, D_2O) 7.68(1H, d, H-6), 6.11(1H, dd, H-1'), 5.71(1H, d, H-5), 4.44(1H, dd, H-3'), 4.49(2H, 2d, H-5'), 3.33(2H, 2d, 4'-C-CH_2), 2.33(2H, m, H-2'), 1.89(3H, s, CH_3CONH); δ_C (75.45MHz, D_2O) 175.60, 167.08, 152.43, 142.86, 102.93, 88.93, 85.57, 71.91, 63.16, 40.37, 39.15, 22.72; ES +ve *m/z* 300$(M+H)^+$, 242$(M-CH_3CONH)^+$.

v) 5-Iodo-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine (15)

Iodine (0.032g, 0.13mmol) was dissolved in 1,4-dioxane (1mL) at ambient temperature. $PhI(COCF_3)_2$ (0.056g, 0.13mmol) and pyridine (0.022g, 0.28mmol,

0.023mL) were then added and the reaction mixture stirred until iodine colour faded. 4'-*C*-(Acetylaminomethyl)-2'-deoxyuridine **(14)** (0.032, 0.11mmol) in pyridine (0.25mL) was then added dropwise to the reaction mixture. The mixture was warmed to 60C for on hour and then allowed to cool to ambient temperature after which the solvents were removed under vacuum. The residue was re-dissolved in 9:1 dichloromethane and then eluted through a short flash silica gel column using 9:1 dichloromethane: methanol. The title compound **(15)** was obtained, with some unreacted starting material, as an amorphous solid. δ_H (300MHz, CD_3OD) 8.52(1H, s, H-6), 6.21(1H, dd, H-1'), 4.48(1H, dd, H-3'), 3.60(2H, 2d, H-5'), 3.39 – 3.29(2H, 2d, 4'-*C*-CH_2), 2.50 – 2.24(2H, m, H-2'), 1.98(3H, s, CH_3CONH).

vi) 5-(*N*-Trifluoroacetyl)propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine (16)

5-Iodo-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine **(15)** (0.1mmol) and CuI (0.01g, 0.05mmol) were weighed into an round –bottom flask equipped with a magnetic stirrer bar. The flask was then fitted with a septum and flushed with dry argon. *N,N*-dimethylformamide (0.5mL) was added to the flask and the contents stirred until all solids were dissolved. A solution of *N*-(trifluoroacetyl)propargylamine (0.05g, 0.33mmol) in triethylamine (0.02g, 0.22mmol, 0.03mL) was then added to the flask and the resulting solution stirred for five minutes. Solid tetrakis(triphenylphosphine) palladium (0.031g, 0.03mmol) was then added and the resulting mixture stirred at ambient temperature for 4 hours. The reaction mixture was then diluted with 1:1 dichloromethane : methanol (2mL) before adding solid sodium hydrogencarbonate. The mixture was stirred for a further hour before filtering through celite. The filtrate was then concentrated under vacuum. The residue was purified by flash column chromatography (95:5 dichloromethane : methanol). Obtained the title compound as an amorphous solid (5.3mg). δ_H (300MHz, CD_3OD) 8.32(1H, s, H-6), 6.22(1H, dd, H-1'), 4.46(1H, dd, H-3'), 4.26(2H, s, propargyl CH_2), 3.59(2H, 2d, H-5'), 3.53(1H, d, 4'-C-CH_2), 3.36(1H, d, 4'-C-CH_2), 2.37(2H, m, H-2'), 1.97(3H, s, CH_3CONH); δ_C (75.45MHz, CD_3OD) 174.33, 164.66, 151.15, 145.79, 124.5 – 110.8(q, CF_3CONH), 99.44, 90.38, 88.35, 86.68, 76.01, 72.66, 64.21, 41.61, 41.24, 22.52, 9.25; ES +ve m/z 470$(M+H)^+$, 493$(M+Na)^+$.

A longer peptidase or amidase cleavable linker is introduced into compound **(16')** (TFA deprotected analogue of compound **(16)**) according to the reaction scheme of Figure 5 in which R is the side chain of any of the natural amino acids.

5-propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine **(16')** is coupled to suitably derivatised (protected) commercially available peptides (e.g Bachem) using standard peptide coupling reagents (e.g. TSTU, EDCI, DCC, HOBt etc.) or assembled by the sequential addition of suitably derivatised (protected) amino acids following standard peptide synthesis protocols either in solution or on solid phase (see, for example, "Synthetic Peptides - A Users Guide" Ed. G. A. Grant; 1992). The resulting nucleoside linker conjugate is converted to the corresponding triphosphate by using the conditions described for the preparation of compound **(6)**. Labelling of the triphosphate so produced is achieved by following the procedure described below for the preparation of compound **(18)**.

Nucleoside analogue **(16)** was converted to a nucleotide and a fluorescent label incorporated according to the reaction scheme of Figure 6.

i) 5-(*N*-Trifluoroacetyl)propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine triphosphate (17)

5-(*N*-Trifluoroacetyl)propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine **(16)** (5.3mg, 0.012mmol) was treated with phosphorous oxychloride (0.04g, 0.024mmol, 2.2uL), tributylammonium pyrophosphate (0.25mL of 0.5M solution) and tributylamine (0.019mL) in trimethylphosphate (0.25mL) and *N,N*-dimethylformamide (0.25mL) according to the procedure used to prepare compound **(6)**.

ii) 5-[*N*-Fluorescein-5(and-6-)carboxamidohexanoyl]propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine triphosphate (18)

5-(*N*-Trifluoroacetyl)propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine triphosphate **(17)** was dissolved in anhydrous DMSO at ambient temperature. 6-[Fluorescein-5(and-6)-carboxamidohexanoic acid N-hydroxysuccinamidyl ester and triethylamine were then added and the reaction mixture stirred at ambient temperature for 18 hours. Reaction mixture was then diluted with water and lyophilised for 48

hours. The crude material was then re-dissolved in 0.1M triethylammonium bicarbonate (1mL) and subjected to preparative reverse phase HPLC purification using 0.1 triethylammonium bicarbonate / water as mobile phase.

Example 4

Synthesis of nucleoside analogues 4'-*C*-(N-Acetylglycylaminomethyl)thymidine **(19)**; 4'-*C*-(*N*-Leucylaminomethyl)thymidine **(21)** and 4'-*C*-(*N*-Glycylaminomethyl)thymidine **(22)** which have a blocking group attached through an amidase cleavable linker at the 4' position.

Figure 7 shows a reaction scheme for synthesis of nucleoside analogues (compounds **(19)** and **(21)**) which have a blocking group attached through an amidase cleavable linker at the 4' position.

i) 4'-*C*-(N-Acetylglycylaminomethyl)thymidine (19)

4'-*C*-(Aminomethyl)thymidine **(4)** (0.025g, 0.1mmol), *N*-acetylglycine (0.023g, 0.2mmol) and TSTU (0.09g, 0.3mmol) were dissolved in anhydrous *N,N*-dimethylformamide (1mL) at ambient temperature. Diisopropylethylamine (0.065g, 0.5mmol, 0.087mL) was then added and the reaction mixture stirred for 18 hours at ambient temperature. The solvent was then removed under vacuum and the residue re-dissolved in 9:1 dichloromethane : methanol and eluted through a flash silica gel column with 9:1 dichloromethane : methanol. The title compound **(19)** was obtained as an amorphous solid on removal of solvent from the appropriate fractions (0.023g). δ_H (300MHz, CD_3OD) 7.78(1H, s, H-6), 6.26(1H, dd, H-1'), 4.48(1H, dd, H-3'), 3.83(2H, s, glycyl CH_2), 3.60(2H, 2d, H-5'), 3.58(1H, d, 4'-C-CH_2), 3.42(1H, d, 4'-C-CH_2), 2.33(2H, m, H-2'), 2.00(3H, s, CH_3CO), 1.86(3H, s, 5-CH_3); ES +ve *m/z* 393$(M+Na)^+$.

ii) 4'-*C*-[N-(9-Fluorenylmethyloxycarbonyl)leucylaminomethyl)thymidine (20)

4'-*C*-(Aminomethyl)thymidine **(4)** (0.025g, 0.1mmol), *N*-(9-fluorenylmethyloxycarbonyl)leucine (0.071g, 0.2mmol), TSTU (0.09g, 0.3mmol) and diisopropylethylamine (0.065g, 0.5mmol, 0.087mL) were combined in DMF (1mL)

according to the procedure used to prepare compound **(19)**. The title compound **(20)** was obtained as an amorphous solid (0.056g). δ_H (300MHz, CD3OD) 7.86 – 7.37(8H, m, Fmoc), 7.74(1H, s, H-6), 6.24(1H, dd, H-1'), 4.56(1H, dd, H-3'), 3.68(2H, 2d, H-5'), 3.40(2H, 2d, 4'-*C*-CH_2), 2.37(2H, m, H-2'), 1.70(3H, m, leucyl CH, CH_2), 1.66(3H, s, 5-CH_3), 1.05(3H, d, leucyl CH_3), 1.02(3H, d, leucyl CH_3); ES +ve *m/z* 607$(M+H)^+$, 629$(M+Na)^+$.

iii) 4'-*C*-(*N*-Leucylaminomethyl)thymidine (21)

4'-*C*-[*N*-(9-Fluorenylmethyloxycarbonyl)leucylaminomethyl)thymidine **(20)** was dissolved in N,N-dimethylformamide. Piperidine was then added and the solution stirred at ambient temperature for 6 hours. The solvent was then removed under vacuum and the residue re-dissolved in water. The aqueous solution was then washed with toluene and 40-60 petroleum ether, and then separated and lyophilised. The lyophilised product was further purified by flash column chromatography (85:15 dcihloromethane : methanol) to give the title compound **(21)** as a resin. δ_C (300MHz, D_2O) 7.47(1H, s, H-6), 6.13(1H, dd, H-1'), 4.65(1H, dd, H-3'), 3.34(2H, 2d, H-5'), 3.28(1H, d, 4'-*C*-CH_2), 3.00(1H, d, 4'-*C*-CH_2), 2.50(1H, dd, leucyl CH), 2.33(2H, m, H-2'), 1.73(3H, s, 5-CH_3), 1.11(2H, m, leucyl CH_2), 0.76(3H, d, leucyl CH_3), 0.71(3H, d, leucyl CH_3); ES +ve *m/z* 385$(M+H)^+$.

The nucleotide equivalent of compound **(19)** is prepared according to the methods for phosphorylating compounds **(5)** or **(7)** (see above).

The nucleoside equivalent of compound **(21)** is prepared by phosphorylation of compound **(20)** according to the methods for phosphorylating compound **(7)** (see above) prior to the deprotection reaction of compound **(20)** to form compound **(21)** as described above.

Figure 8 shows a reaction scheme for synthesising a further nucleoside having an amidase cleavage site at the 4' position (compound **(22)**).

i) 4'-*C*-(*N*-Glycylaminomethyl)thymidine (22)

4'-*C*-(*N*-trifluoroacetylglycylaminomethyl)thymidine **(5)** (17mg, 0.04mmol) was dissolved in concentrated aqueous ammonia solution (1mL) and allowed to stand overnight at ambient temperature. The reaction mixture was then lyophilised to give the title compound **(22)** as a colourless resin (0.0088g, 67%). δ_H (300MHz, D_2O) 7.46(1H, s, H-6), 6.14(1H, dd, H-1'), 4.48(1H, dd, H-3'), 3.70(2H, s, glycyl CH_2), 3.47(2H, 2d, H-5'), 3.40(2H, 2d, 4'-*C*-CH_2), 2.35(2H, m, H-2'), 1.73(3H, s, 5-CH_3); ES +ve *m/z* 329$(M+H)^+$.

A nucleosides having formula **(22)** is converted to the corresponding triphosphate by using the conditions described for the preparation of compound **(8)**.

Figure 9 shows a reaction scheme for synthesising a nucleotide having a penicillin amidase cleavage site at the 4' position (compound **(31)**).

i) *N*-{α-[3'-*O*-((*tert*butyl)dimethylsilyl)-5'-*O*-((*tert*butyl)diphenylsilyl)-4'-methyloxythymidyl]phenyl}phenylacetamide (29).

3'-*O*-[(*tert*butyl)dimethylsilyl]-5'-*O*-[(*tert*butyl)diphenylsilyl]-4'-hydroxymethyl thymidine (Marx et. al. *Helv. Chim. Acta*. 1996, **79**, *1980 – 1994* and references cited therein) **(28)** is treated with *N*-[α-thioethylphenyl]phenyl acetamide **(24)** and *N*-iodosuccinimide in accordance with the procedure described for the preparation of compound **(25)** to obtain the title compound **(29)**.

ii) *N*-{α-[4'-methyloxythymidyl]phenyl}phenylacetamide (30).

N-{α-[3'-*O*-((*tert*butyl)dimethylsilyl)-5'-*O*-((*tert*butyl)diphenylsilyl)-4'-methyloxythymidyl]phenyl}phenylacetamide **(29)** is treated with tetrabutylammonium fluoride in tetrahydrofuran in accordance with the procedure used to prepare compound **(26)** to obtain the title compound **(30)**.

iii) *N*-{α-[4'-methyloxythymidyl]phenyl}phenylacetamide triphosphate (31).

N-{α-[4'-methyloxythymidyl]phenyl}phenylacetamide (30) is treated with 2-chloro-4H-1,3,2-dioxaphosphorin-4-one, pyridine, tributylammonium pyrophosphate, iodine solution and tributylamine in accordance with the procedure used to prepare compound **(27)** to obtain the title compound **(31)**.

Labelling of the triphosphates so produced is achieved by following the procedure described for the preparation of compound **(18)**.

Example 5

Synthesis of a nucleotide with an enzyme (penicillin amidase) cleavable blocking group at the 3'-position.

Figure 10 shows a reaction scheme for the synthesis of a nucleotide of formula **(27)**.

i) *N*-[α-(5'-*O*-*tert*butyldiphenylsilyl-3'-*O*-thymidyl)-phenyl)]-phenylacetamide (25)

5'-*O*-tertbutyldimethylsilyl thymidine (0.5g, 1mmol) **(23)** and *N*-[α-thioethylphenyl]phenyl acetamide (Flitsch et. al. *Tetrahedron Letters* 1998, **39**, *3819 – 3822* and references cited therein; Flitsch et. al. PCT WO 97/20855) **(24)** (0.28g, 1mmol) were dissolved in anhydrous dichloromethane (5mL) at ambient temperature. Crushed, activated 4A molecular sieves (1g) were then added and the mixture stirred for 15 minutes at ambient temperature before cooling to 0C on an ice bath. *N*-Iodosuccinimide (0.33g, 1.5mmol) was then added as a solid to the cooled solution. The resulting solution was then stirred for 3 hours at 0C. Saturated sodium thiosulfate solution (10mL) was then added and the organic layer separated. The aqueous layer was back extracted with a portion of dichloromethane (10mL). The combined organic extracts were then dried over magnesium sulfate, filtered and then concentrated under vacuum. Purification by flash column chromatography (1:1 40-60 petrol ether: ethyl acetate) gave the title compound (1:1 mixture of diastereoisomers) as a white foam on removal of solvent from the appropriate fractions. Yield= 0.4g (55%). δ (300MHz, d_6-DMSO) 11.35(1H, s, N^3-H), 9.05(1H, 2d, amide N-H), 7.59 – 7.15(22H, m, Ph_2Si, Ph, $PhCH_2CO$, α-H, H-6), 6.23 – 6.14(1H, m, H-1'), 4.46 – 3.29(4H, m, H-5', H-4', H-3'), 2.48 – 2.17(2H, m, H-2'), 1.44, 1.41(3H, 2s, 5-CH_3), 0.98, 0.95(9H, 2s, *t*BuSi); δ (75.45MHz, d_6-DMSO) 170.87, 163.54, 150.38, 139.21, 135.85, 135.23, 134.88, 132.85, 132.26, 130.02, 129.04, 128.28, 128.24, 127.97, 126.30, 109.74, 84.40, 83.91, 83.67, 83.47, 79.03, 78.60, 76.54, 76.40, 63.97, 59.72, 55.33, 48.57, 42.23, 37.45, 26.60, 20.74, 18.83, 14.06, 11.72.

ii) *N*-[α-(3'-*O*-thymidyl)-phenyl)]-phenylacetamide (26)

N-[α-(5'-*O*-*tert*butyldiphenylsilyl-3'-*O*-thymidyl)-phenyl)]-phenylacetamide **(25)** (0.4g, 0.56mmol) was dissolved in tetrahydrofuran (10mL) at ambient temperature. A 1M solution of tetrabutylammonium fluoride in tetrahydrofuran (0.56mL) was then added to the solution of the nucleoside. After stirring for 4 hours methanol (2mL) was added and the solvent removed under vacuum to give a white foam. Flash column chromatography (95:5 dichloromethane: methanol) gave the title compound as a white foam (mixture of diastereoisomers). Yield= 0.22g (85%). δ (300MHz, d_6-DMSO) 11.30(1H, s, N^3-H), 9.03(1H, 2d, amide N-H), 7.67, 7.63(1H, 2s, H-6), 7.43 – 7.18(11H, m, Ph, $PhCH_2CO$, α-H), 6.19 – 6.11(1H, m, H-1'), 5.07(1H, t, br, 5'-OH), 4.28 – 3.98(3H, m, H-4', H-3'), 3.56(2H, m, H-5'), 2.40 – 2.20(2H, m, H-2'), 1.75, 1.74(3H, 2s, 5-CH_3); δ (75.45MHz, d_6-DMSO) 170.84, 170.78, 139.44, 135.94, 135.88, 129.02, 128.28, 128.26, 128.23, 128.15, 126.43, 126.35, 109.49, 84.65, 83.75, 78.85, 77.20, 76.95, 59.72, 42.20, 20.74, 14.06, 12.24.

iii) *N*-[α-(3'-*O*-thymidyl)-phenyl)]-phenylacetamide triphosphate (27)

N-[α-(3'-*O*-thymidyl)-phenyl)]-phenylacetamide **(26)** (0.11g, 0.24mmol), 2-chloro-4H-1,3,2-dioxaphosphorin-4-one (0.24mL of 1M solution in 1,4-dioxane), pyridine (0.24mL), tributylammonium pyrophosphate (0.72mL of 1M solution in DMF), iodine (0.16g dissolved in 7.2mL pyridine / 0.14mL water) and tributylamine were combined according to the procedure used to prepare compound **(8)**.

<u>Example 6</u>

<u>Preparation of a nucleotide with a penicillin amidase cleavable linker and penicillin amidase cleavable 3'-blocking group.</u>

Figure 11 shows a reaction scheme for the synthesis of a compound of formula **(36)**.

i) 5-Hydroxymethyl-5',3'-di-*O*-*p*-toluyl-2'-deoxyuridine **(32)** (prepared using well established procedures described, for example, in Chemistry of Nucleosides and Nucleotides; Volume 1. Ed. L. B. Townsend, 1988) and *N*-[α-thioethyl-*N*'-

trifluoroacetylaminopropyl benzamide] phenylacetamide **(33)** (prepared from procedures in Katritzky et al.; Synthesis, 1993, 445 – 456) is combined in the presence of *N*-iodosuccinimide according to the procedure used to obtain compound **(25)** to afford the intermediate **(34)**.

ii) Compound **(34)** is converted to the intermediate **(35)** by treatment with sodium methoxide in methanol, followed by ethyl trifluoroacetate in methanol.

iii) Compound **(35)** may then be converted to nucleoside **(36)** using the procedure outlined for the preparation of compound **(26)**. Conversion to a triphosphate is achieved by using the procedure described for the preparation of compound **(27)**. Labelling of the triphosphate so produced is achieved by following the procedure described for the preparation of compound **(18)**.

<u>Example 7</u>

<u>Preparation of a nucleoside analogue containing a fluorescent group attached via an enzyme cleavable linker.</u>

Figure 12 shows a reaction scheme for the synthesis of a compound of formula **(42)**.

i) 5-iodo-2'-deoxyuridine **(37)** was reacted with TFApropargylamine in a palladium ($Pd(PPh_3)_4$) catalysed coupling reaction in the presence of CuI and DMF to introduce the trifluoroacetylpropargylamino arm.

ii) 5-propargylamino-2'-deoxyuridine **(39)** was then obtained by stirring with concentrated aqueous ammonia

iii) Coupling the tetrapeptide *N*-Boc-Ala-Gly-Gly-Gly-OH (Bachem Ltd, UK) (tBOC-NH-AGGG-OH) to **(39)** was achieved by treatment with EDCI.HCl and *N*-hydroxysuccinimide in dimethylformamide.

iv) The *N*-terminal Boc group was then removed by treatment with 50% trifluoroacetic acid in dichloromethane to give the peptide-nucleoside conjugate **(41)**.

v) Labelling of the tetrapeptide N-terminus was achieved by treating **(41)** with a molar excess of fluorescein hexanoic acid NHS ester (6-(Fluorescein-5(and-6) carboxamidohexanoic acid NHS ester) in triethylamine / dimethylformamide. The product **(42)** was purified by thin layer chromatography.

Example 8

Enzyme cleavage reactions

a) Cleavage of the 4' blocking group from compound **(22)**

1mg of 4'-*C*-(*N*-Glycylaminomethyl)thymidine **(22)** or 4'-*C*-(N-Acetylglycylaminomethyl)thymidine **(19)** is incubated at 37°C in 60mM sodium phosphate buffer pH 7.0 with 0.5 units aminopeptidase M (Calbiochem 164598), to a final volume of 200ul.

b) Cleavage of 4' blocking group from compound **(21)**

1mg of 4'-*C*-(*N*-Leucylaminomethyl)thymidine **(21)** is incubated at 37°C in 47mM sodium phosphate pH 7.2 with 3.3% methanol and 0.25 units leucine aminopeptidase (Sigma L0632), to a final volume of 200ul.

c) Cleavage of 3' blocking group from compound **(26)**

1mg of *N*-[α-(3'-*O*-thymidyl)-phenyl)]-phenylacetamide **(26)** is incubated at 37 °C in 50mM potassium phosphate pH 7.5 with 2 units of penicillin amidase (Fluka P3319), to a final volume of 200ul.

The products resulting from a), b) or c) above are analysed by thin layer chromatography using reverse phase plates with 10% acetonitrile in water or by electrospray mass spectrometry. Successful cleavage reactions yield nucleosides running concurrently with 4'-aminomethylthymidine which is used as a standard.

Mass spectrometric detection of the thymidine daughter ion (mass 242) arising from fragmentation of 4'-*C*-(*N*-aminomethyl)thymidine suggested that enzyme induced cleavage of the amino acid from the leucyl-containing compound (4'-*C*-(*N*-Leucylaminomethyl)thymidine **(21)**) had occurred.

d) Cleavage of the fluorophore from the nucleoside of compound **(42)**

1mg of the substrate **(42)** was dissolved in a buffer containing 0.1M sodium phosphate, pH 6.3, 5mM EDTA, 60μM β-Mercaptomethanol. Enzyme digestion was started by adding 1unit of papain (150 units/mg, Europa Ltd) in a final reaction volume of 200 μl followed by incubation at 50°C for 2h. The reactions were spotted on a thin layer chromatography plate and air dried before chromatography for 20 min. in a 4:1 mixture of dichloromethane/methanol.

Following thin layer chromatography, the undigested substrate control migrated from the origin with no ultraviolet light absorbing material present at the origin and a clear fluorescence observable proximal to the solvent front. In contrast the non-fluoresceinated nucleoside control appeared as a strongly ultraviolet light absorbing material that showed little or no migration from the origin. As evidenced by thin layer chromatography, the papain digested material resulted in the production of a non-fluorescent, strongly ultraviolet light absorbing material at the origin that was consistent with the cleavage of the linker and separation of the nucleoside from the fluor.

Example 9

DNA polymerase assays

DNA polymerase assays were performed using Sequence 1 as primer and Sequence 2 as template.

Sequence 1 5'[Cy3]TAACTCATTAACAGGATC 3'

Sequence 2 5'AT TCG CGG TAT TCT GGT ATG AAG CTT TTA GAT CCT GTT AAT GAG TTA GTA3'

The template (Sequence 2) was designed such that the base (underlined) immediately adjacent to the 3' end of the hybridised primer is complementary to the nucleotide test compound. Nucleotide test compounds were compounds **(6)** (4'-C-(glycylaminomethyl)thymidine triphosphate), **(8)** (4'-C-(N-trifluoroacetyaminomethyl)thymidine triphosphate) and **(9)** (4'-C-(aminomethyl)thymidine triphosphate) as described above.

Primer extension reactions were performed using the above primers in the presence of 40 to 80µM of nucleotide triphosphates. Positive control reactions contained all four native nucleotide triphosphates. For the other reactions the TTP was replaced with either a 2:1 ratio of TTP/dTTP for the terminator control, or entirely by the test compound. The enzymes were used at a final concentration of 0.175units/µl and a magnesium ion concentration of 2.5mM in buffers supplied by the manufacturer. The primer and template concentrations were 0.5pmol/µl and 2pmol/µl respectively.

Extension reactions with thermophilic enzymes were performed by initially denaturing the fully assembled reactions at 70°C for 5 min. followed by 45°C for 20-40 minutes. The following enzymes were used: Taq DNA polymerase (Taq); Thermosequenase (TS); Thermosequenase II (TSII); Thermosequenase E (TSE), Tfl DNA polymerase (Tfl); Tth DNA polymerase (Tth), deltaTts DNA polymerase (deltaTts), deltaTts D DNA polymerase (deltaTTSD) (all from APBiotech); Pfu DNA polymerase (Pfu) (Stratagene Ltd); Vent™ DNA polymerase (Vent) (New England Biolabs).

Extension reactions performed with mesophilic polymerases were carried out by mixing all the components of the reaction except the enzyme and heating to 70°C for 5 min. After the reactions had cooled to room temperature 0.175units of the enzyme were added and the reactions heated to 37°C for 20-40 minutes. The following

enzymes were used: T7 DNA polymerase (T7); Sequenase 2™ (Seq); Klenow fragment of DNA polymerase 1 (Klen); phi-29 DNA polymerase (Phi-29); T4 DNA polymerase (T4) (all from AP Biotech) and Bst DNA polymerase (Bst) (Cambio Ltd).

Reverse transcriptase reactions were performed essentially as described above for thermophilic and mesophilic enzymes. The template for these reactions was the RNA equivalent of sequence 2. All the thermophilic enzymes were assayed in Tth reaction buffer (Cambio Ltd, Cambridge UK) containing 2.5mM Manganese ions. The following enzymes were used: Thermosequenase (TS1); Thermosequenase II (TSII); delta Tts DNA polymerase (deltaTts); delta Tts D DNA polymerase (deltaTtsD) (all from APBiotech); Retrotherm reverse transcriptase (Retroth) (Cambio Ltd); Tth DNA polymerase (Tth) (Cambio Ltd.).

Mesophilic reverse transcriptases (AMV RT; MuLV RT; SAV RT and HIV RT all from APBiotech) were assayed in their respective buffers with a final enzyme concentration of 4units/μl.

The reactions were stopped by adding a 0.5 volume aliquot of 80% formamide containing 0.1%(w/v) bromophenol blue and 0.1%(w/v) cresol blue. The reaction products were separated by denaturing gel electrophoresis at a constant power of 45W on 16% polyacrylamide gels containing 50%urea.

Gels were scanned for Cy3 fluorescence on a Molecular Dynamics Fluorimager and the data analysed using Molecular Dynamics ImageQuant 5 software.

The incorporation of a test nucleotide triphosphate results in the appearance on a polyacrylamide gel of a single band corresponding to a single base extension product. If, however, the enzyme adds additional bases then further bands would be observed that corresponded to either the full-length product or positions of other complementary bases in the template.

The results were graded as follows:

-	No extension within assay time
+/-	up to 25% of primer extended by at least one base within assay time
+	up to 50% of primer extended by at least one base within assay time
++	up to 75% of primer extended by at least one base within assay time
+++	essentially all primer extended by at least one base within assay time
T	Incorporation results in termination
E	Incorporation permits further chain extension
T/E	Predominant product is primer extended by one base, but evidence of some further extension also observed.

Table 1 shows the results of **Thermostable DNA polymerase reactions**

	compound (6)		compound (8)		compound (9)	
Taq	+/-	T	+/-	T	+/-	T
TS	+	T	+	T	+	T
TSII	+	T	+	T	+	T
Pfu	-		-		-	
Vent	+/-	T	+/-	T	+/-	T
TSE	++	T	++	T	++	T
Tfl	+/-	T	+/-	T	+/-	T
Tth	+	T	++	T	++	T
deltaTts	+	T	+	T	+T	T
deltaTTSD	++	T	++	T/E	++	T/E

Table 2 shows the results of **Mesophilic DNA polymerase reactions**

	compound (6)		compound (8)		compound (9)	
T7	+/-	T/E	+/-	T/E	+/-	T/E
Bst	+	T	+	T	+	T
Seq	+	T	+	T	+	T
Klen	-		-		-	
phi-29	+/-	T	+/-	T	+/-	T
T4	+	T	+	T	+	T

Table 3 shows the results of reactions with **Thermostable polymerases under reverse transcription conditions**

	compound (6)		compound (8)		compound (9)	
TS I	+	T/E	+	T/E	+	T/E
TSII	+	E	+	E	++	E
Retroth.	++	E	++	E	+++	E
Tth	+++	T	+++	T	+++	T
deltaTts	+++	T	+++	T	+++	T
deltaTtsD	+++	E	+++	E	+++	E

Table 4 shows the results of reactions with **Mesophilic reverse transcriptase.**

	compound (6)		compound (8)		compound (9)	
AMV RT	-		-		-	
MuLV RT	++	E	++	E	++	E
SAV RT	+	E?	++	E	++	E
HIV RT	+++	E	+++	E	+++	E

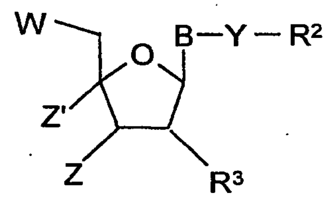
Summary

Compounds **(6)** (4'-C-(glycylaminomethyl)thymidine triphosphate), **(8)** (4'-C-(trifluoroacetylaminomethyl)-thymidine triphosphate) and **(9)** (4'-C-(aminomethyl)thymidine triphosphate) were effective substrates for a range of polymerases and that their incorporation resulted in termination. When reverse transcriptase conditions were used, some extension was observed suggesting that a larger blocking group may be required for those nucleotides when used with certain enzymes under certain specific conditions.

CLAIMS:

1. A nucleotide comprising a reporter moiety and a polymerase enzyme blocking moiety characterised in that the reporter moiety does not also act as a polymerase enzyme blocking moiety.

2. A compound of Formula I



(I)

wherein W is a phosphate group
B is a base
Y is a linker comprising an enzyme-cleavable group
R^2 is a reporter moiety
R^3 is selected from H or OH
Z and Z' are selected from H, OH, or a group X-R^1, wherein X is a linker comprising an enzyme-cleavable group and R^1 is a polymerase enzyme blocking group, provided that at least one of Z and Z' is X- R^1.

3. A compound as claimed in claim 2 wherein W is a triphosphate.

4. A compound as claimed in any of claims 2 or 3 wherein B is selected from the bases A, C, G and T or their analogues.

5. A compound as claimed in any of claims 2 to 4 wherein enzyme-cleavable groups, X and Y, are the same.

6. A compound as claimed in any of claims 2 to 5 wherein X and Y are groups cleavable by enzymes selected from esterases, phosphatases, peptidases, amidases, glycosidases or phosphorylases.

7. A compound as claimed in any of claims 2 to 6 wherein X and/or Y is a group cleavable by an amidase.

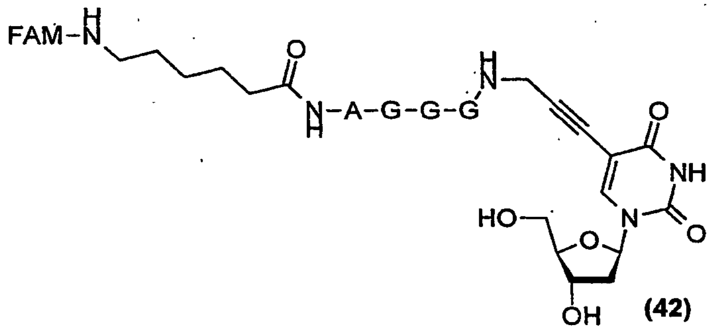
8. A compound as claimed in any of claims 2 to 7 wherein R^2 is a fluorophore, preferably selected from fluoresceins, rhodamines, coumarins, BODIPY™ dyes, cyanine dyes and squarate dyes.

9. A compound as claimed in any of claims 2 to 8 wherein R^1 is selected from CH_3, glycyl or leucyl groups.

10. A compound as claimed in any of claims 2 to 9 wherein Z' is X-R^1 and, preferably, Z is OH.

11. A compound as claimed in any of claims 2 to 9 wherein R^1 is not a reporter moiety.

12. A chemical intermediate selected from the group consisting of: 4'-*C*-(Glycylaminomethyl)thymidine triphosphate; 4'-*C*-(*N*-trifluoroacetylaminomethyl)thymidine triphosphate; 4'-*C*-(Aminomethyl)thymidine triphosphate; 5-(*N*-Trifluoroacetyl)propargylamino-4'-*C*-(acetylaminomethyl)-2'-deoxyuridine; 4'-*C*-(N-Acetylglycylaminomethyl)thymidine; 4'-*C*-(*N*-Leucylaminomethyl)thymidine; 4'-*C*-(*N*-Glycylaminomethyl)thymidine (shown in Figure 8); *N*-{α-[4'-methyloxythymidyl]phenyl}phenylacetamide triphosphate; *N*-[α-(3'-*O*-thymidyl)-phenyl)]-phenylacetamide triphosphate, *N*-{α-[3'-*O*-(5- *N*-(α-methyloxy-*N*'-trifluoroacetylaminopropyl benzamide) phenylacetamide –2'-deoxyuridyl]-phenyl}-phenylacetamide and a compound of formula:



(42)

13. A set of nucleotides characterised in that the set contains at least one compound of Formula I.

14. A set of nucleotides as claimed in claim 13 comprising each of the four natural bases A, G, C and T (or their analogues).

15. A set of nucleotides as claimed in any of claims 13 or 14 further comprising at least two compounds of Formula I having different bases, B, and characterised in that each compound of Formula I has a different reporter moiety, R^2.

16. A set of nucleotides as claimed in any of claims 13 to 15 comprising four compounds of Formula I characterised in that each compound has a different base, B, such that each of the bases A, G, C and T, or analogues thereof, are present and each of the four compounds of Formula I has a reporter moiety which is distinguishable from all the reporter moiety of each of the compounds of Formula I having the other three bases.

17. A method for nucleic acid molecule sequencing comprising the steps of:

a) immobilising a complex of a primer and a template to a solid phase

b) incubating with a polymerase in the presence of a compound of Formula I.

18. A method as claimed in claim 17 further comprising the steps of:

c) detecting the incorporation of a compound of Formula I

d) incubating in the presence of enzyme under suitable conditions for enzymatic cleavage of the enzyme-cleavable groups X and Y.

19. A method as claimed in claim 18 further comprising:
e) repeating steps a)-d).

20. A method as claimed in claim 18 or 19 wherein the enzyme in step d) is an amidase.

FIGURE 1

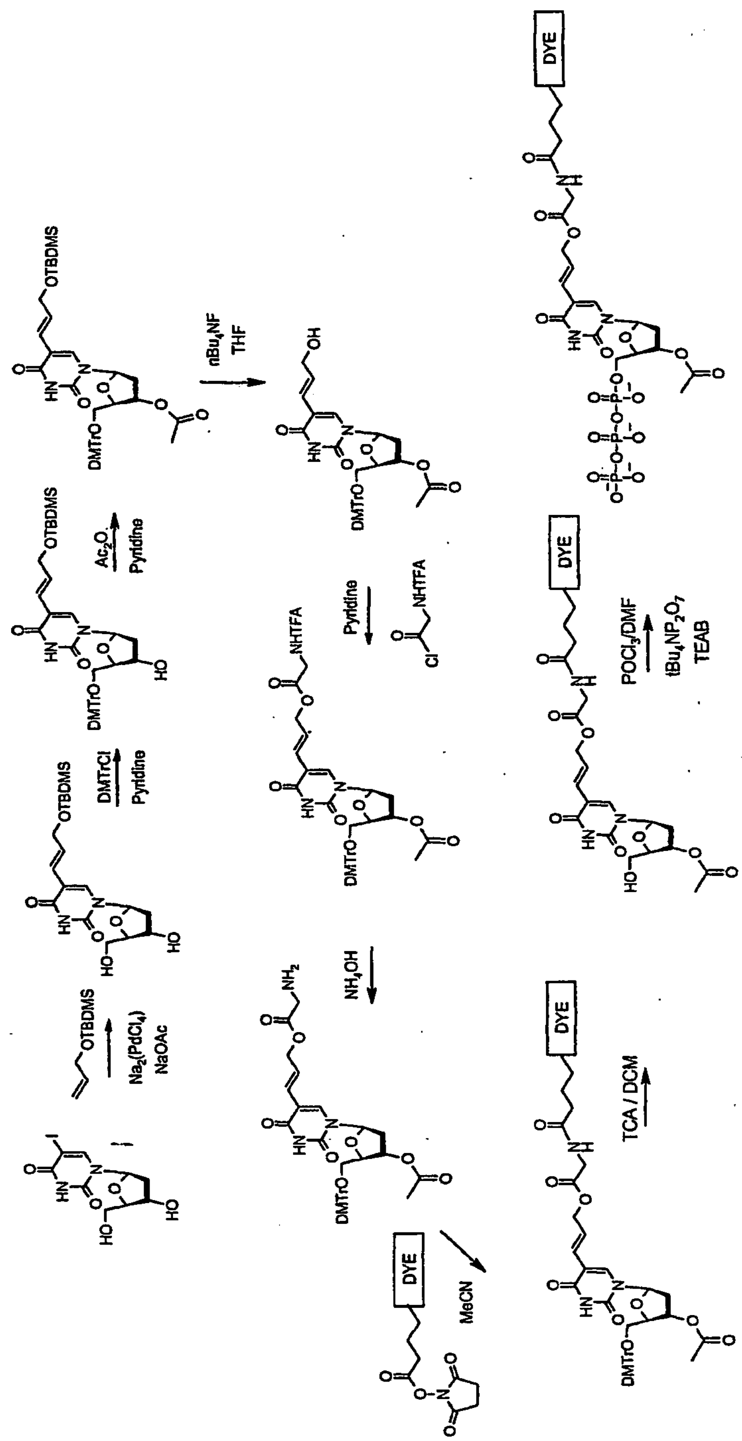
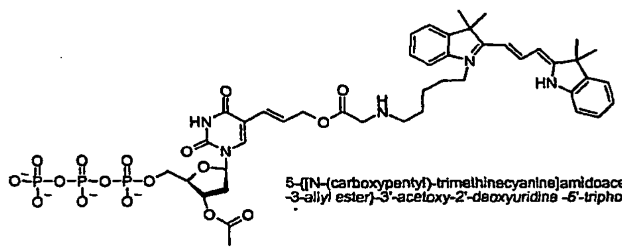
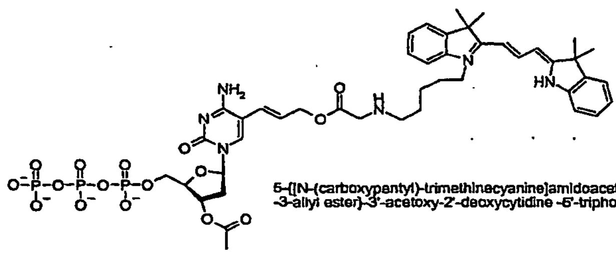

OTBDMS
Na2(PdCl4)
NaOAc
DMTrCl
Pyridine
Ac2O
Pyridine
nBu4NF
THF
NHTFA
Pyridine
NH4OH
DYE
MeCN
TCA / DCM
POCl3/DMF
tBu4NP2O7
TEAB

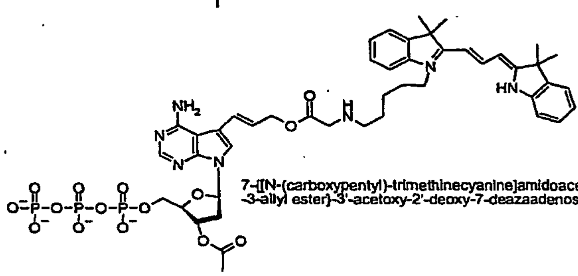
FIGURE 2



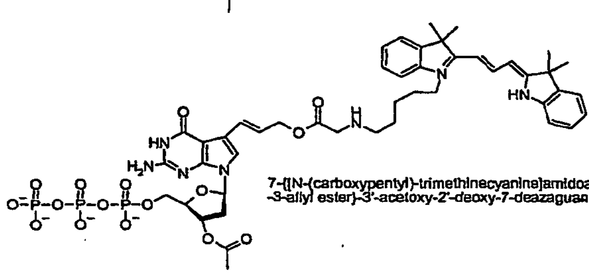
5-{[N-(carboxypentyl)-trimethinecyanine]amidoacetic acid -3-allyl ester}-3'-acetoxy-2'-deoxyuridine -5'-triphosphate



5-{[N-(carboxypentyl)-trimethinecyanine]amidoacetic acid -3-allyl ester}-3'-acetoxy-2'-deoxycytidine -5'-triphosphate



7-{[N-(carboxypentyl)-trimethinecyanine]amidoacetic acid -3-allyl ester}-3'-acetoxy-2'-deoxy-7-deazaadenosine -5'-triphosphate



7-{[N-(carboxypentyl)-trimethinecyanine]amidoacetic acid -3-allyl ester}-3'-acetoxy-2'-deoxy-7-deazaguanosine -5'-triphosphate

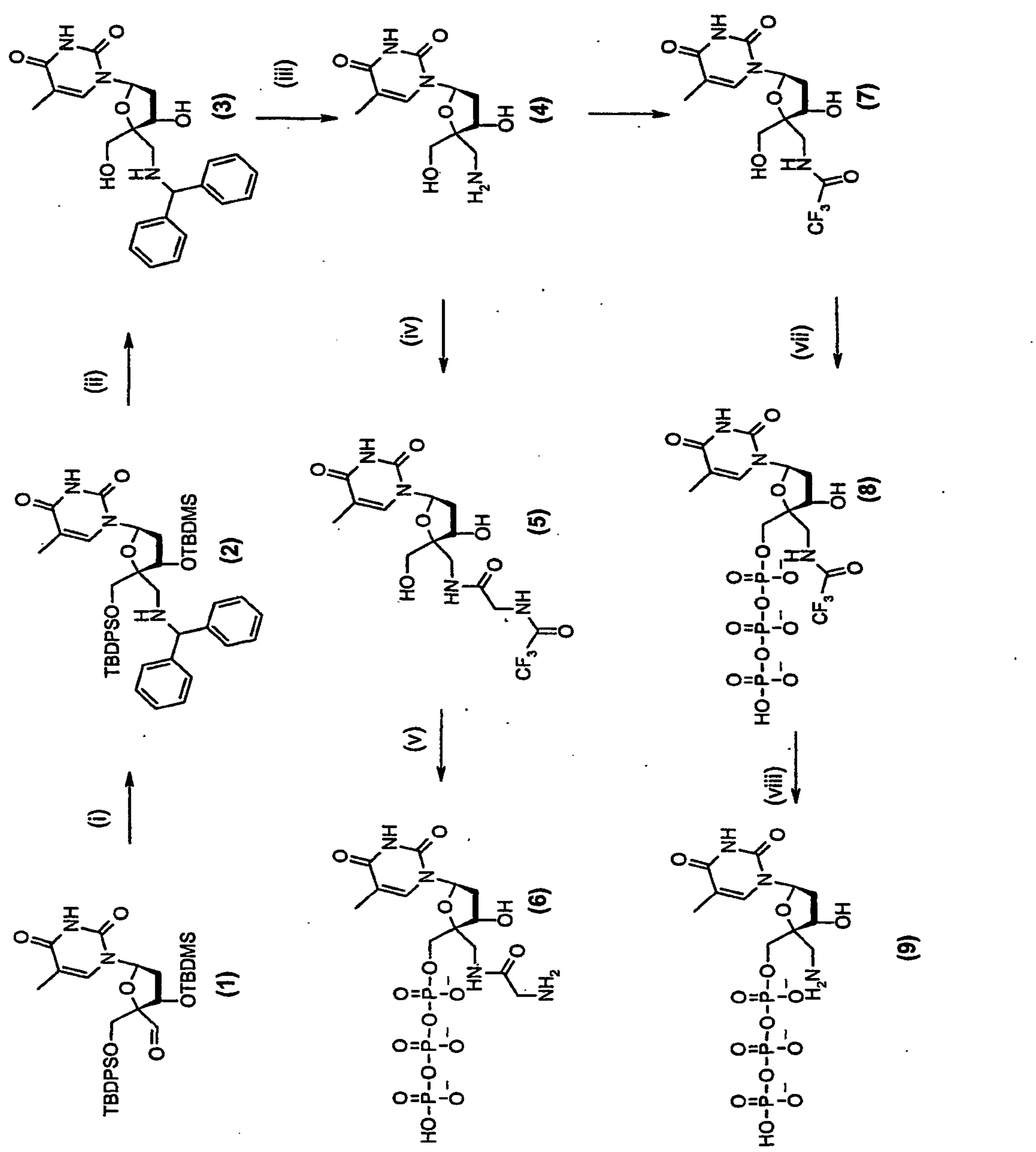


FIGURE 3

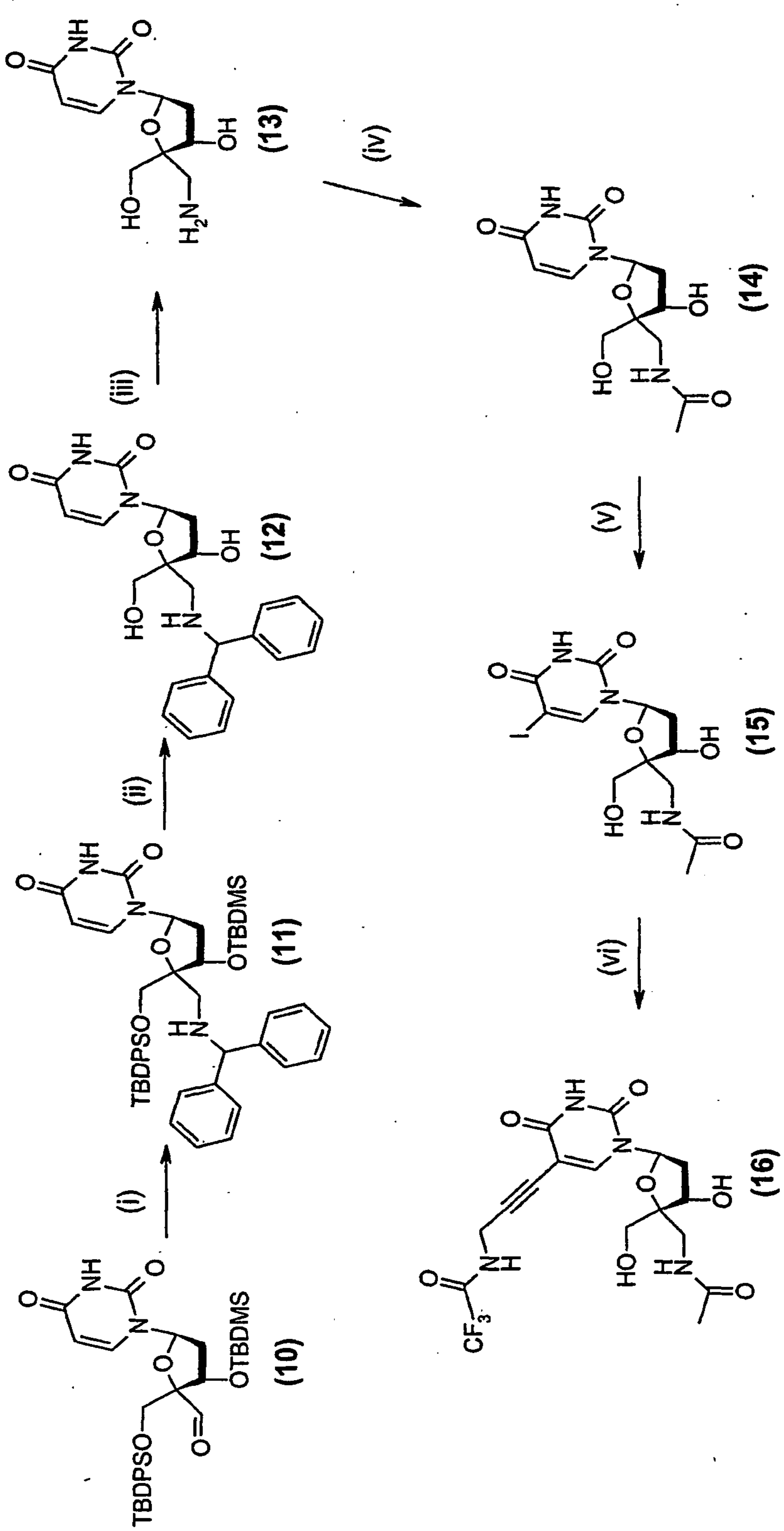
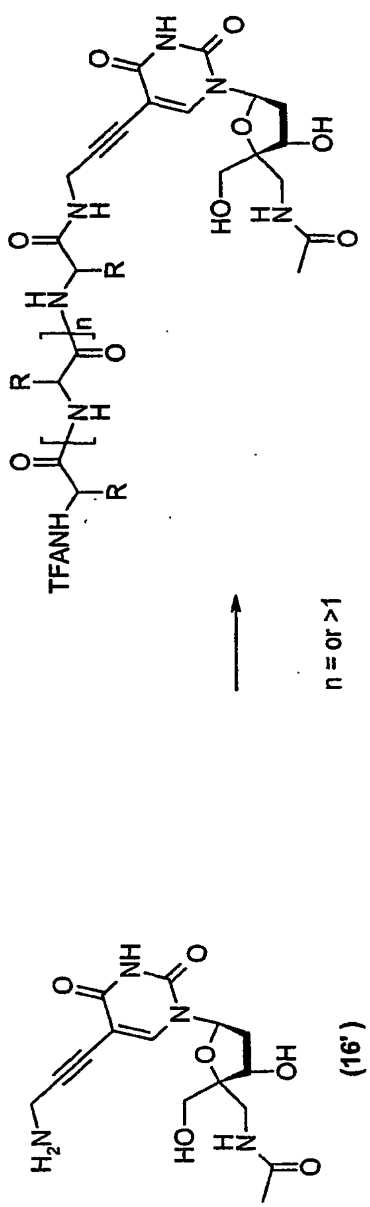

(10)
(11)
(12)
(13)
(14)
(15)
(16)
(i)
(ii)
(iii)
(iv)
(v)
(vi)
TBDPSO
OTBDMS
HO
OH
NH
O
N
H
H_2N
CF_3
I

FIGURE 4

FIGURE 5



(16')

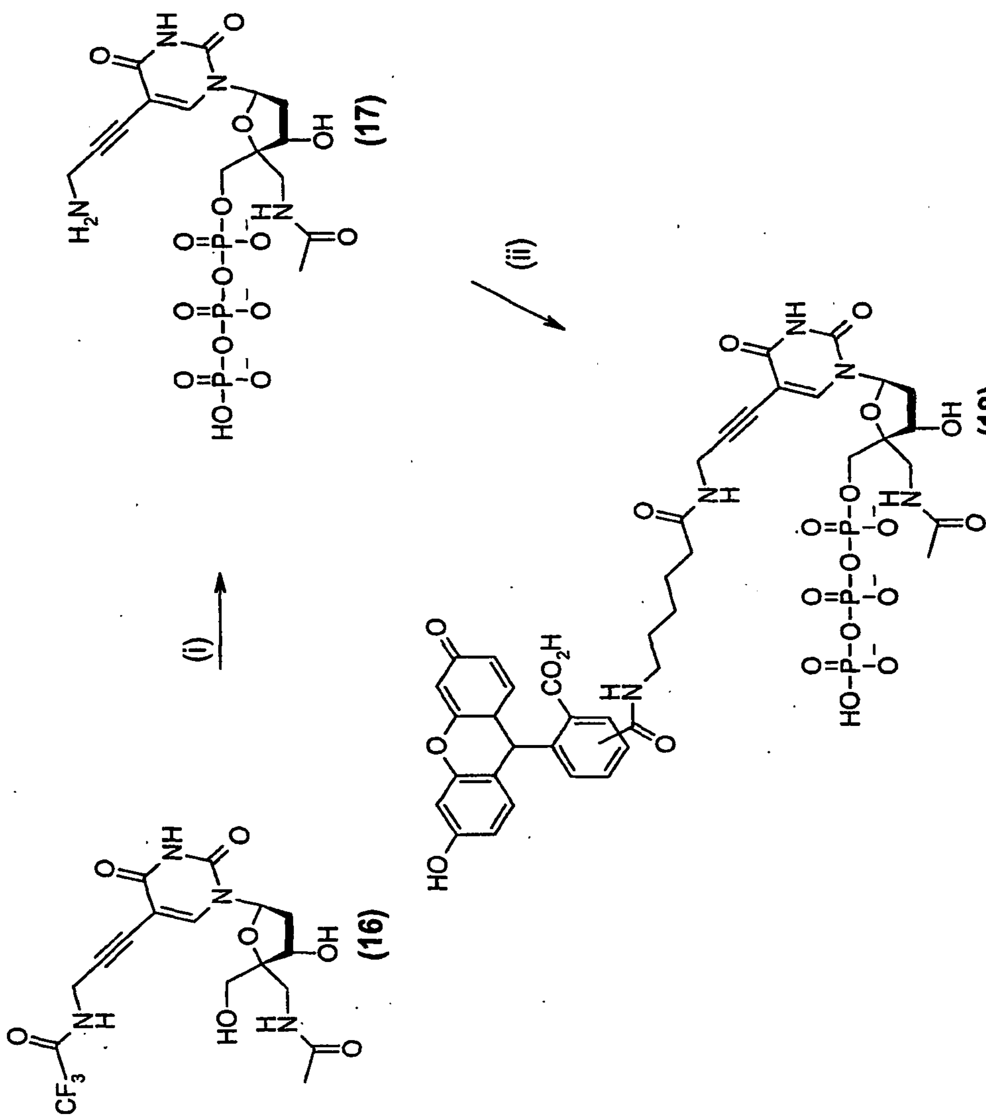


FIGURE 6

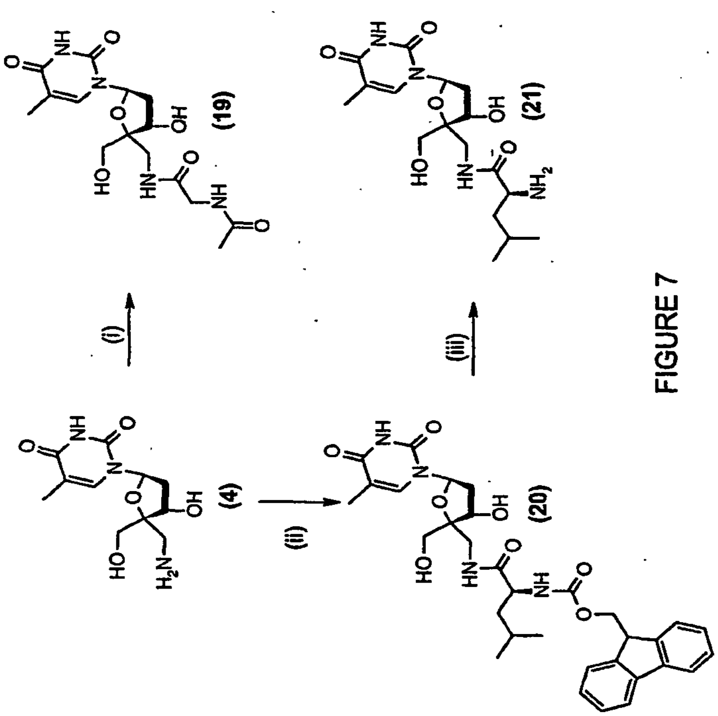


FIGURE 7

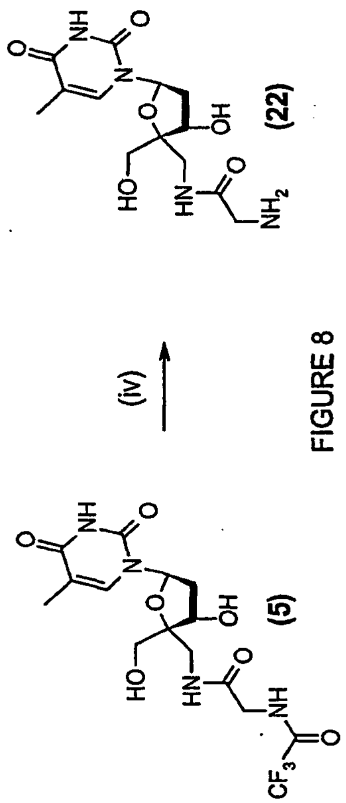


FIGURE 8

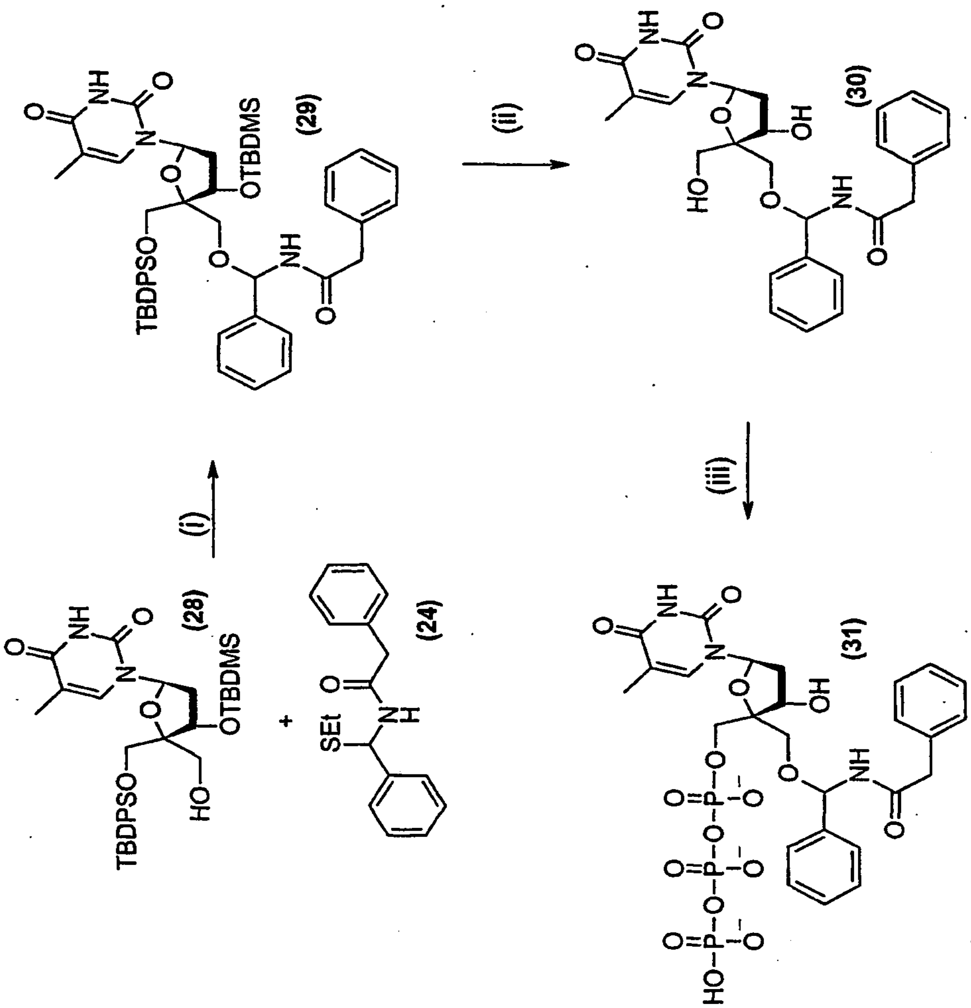


FIGURE 9

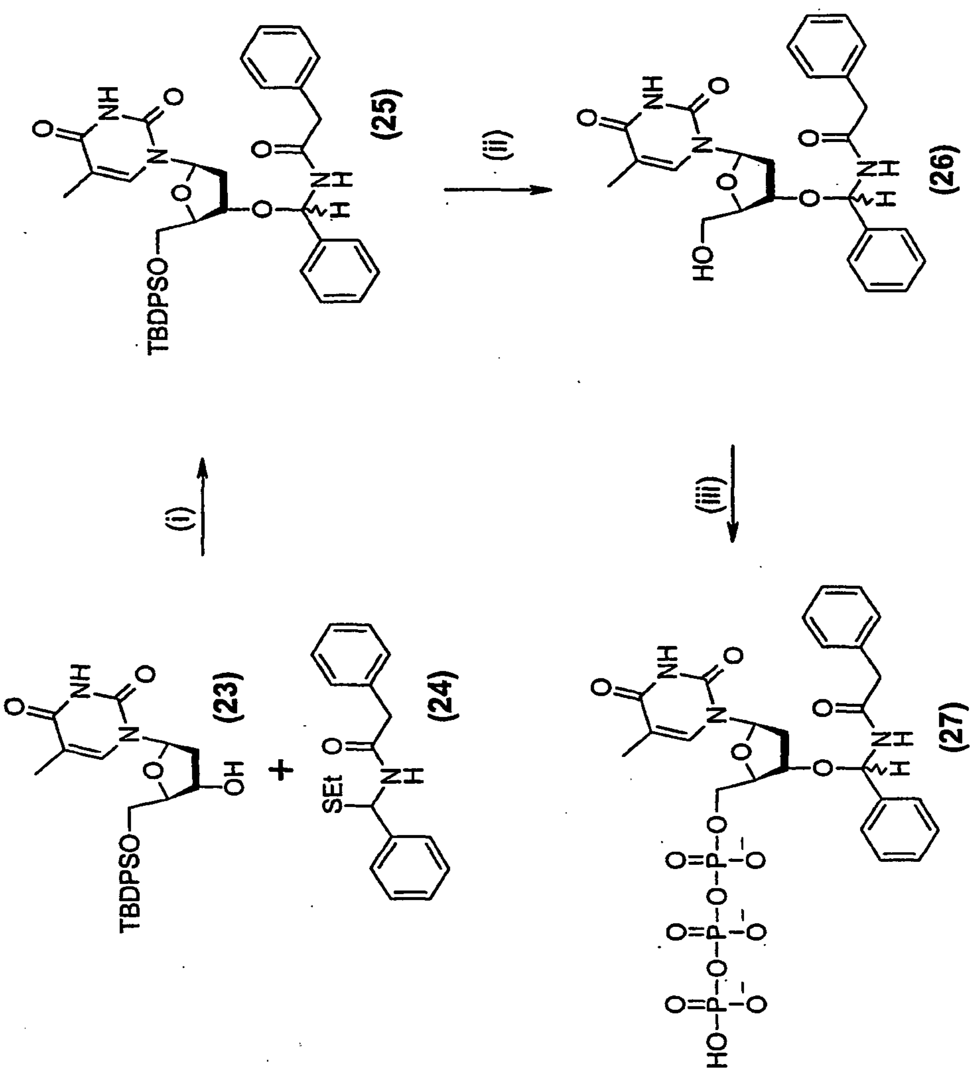


FIGURE 10

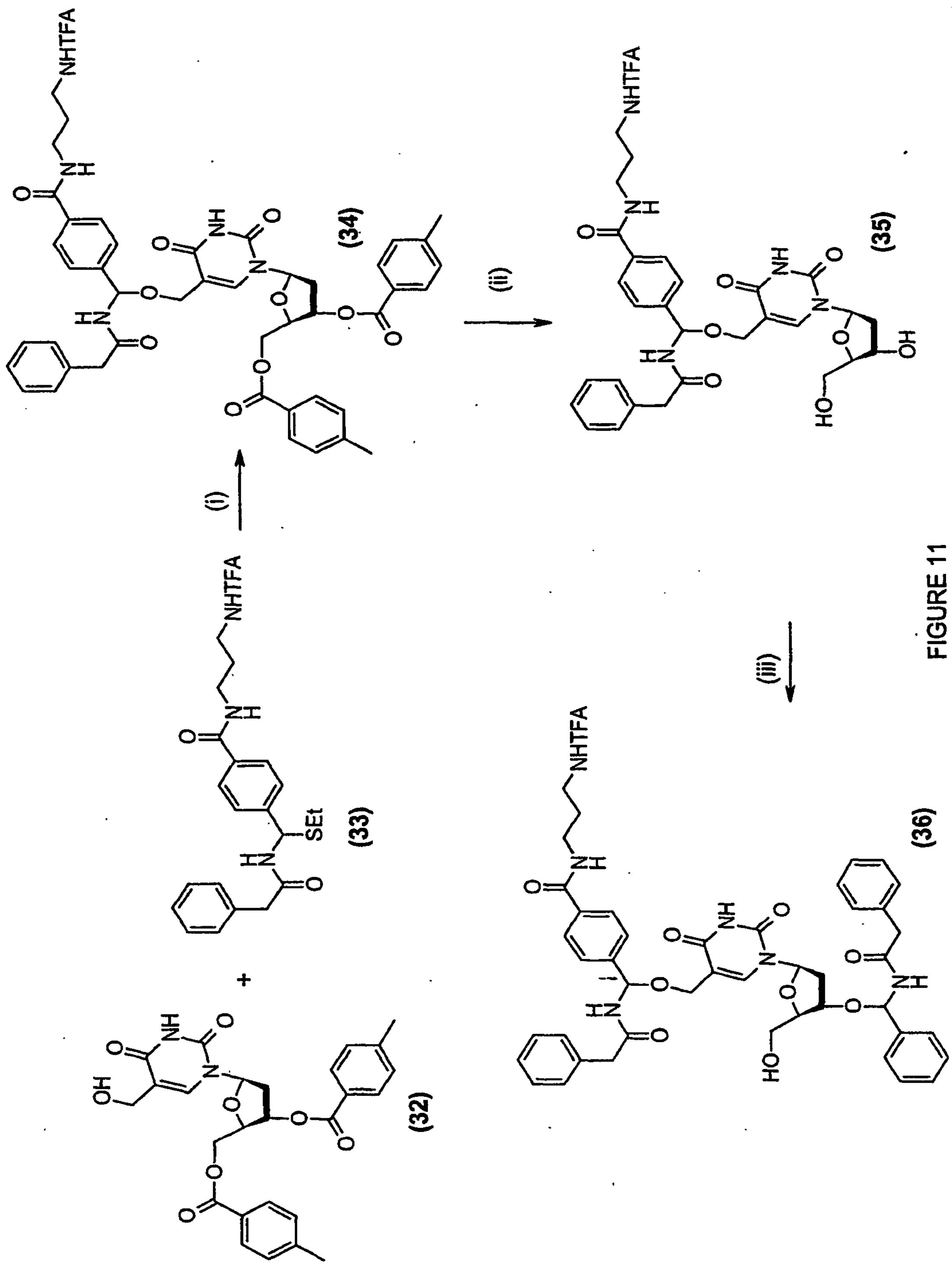


FIGURE 11

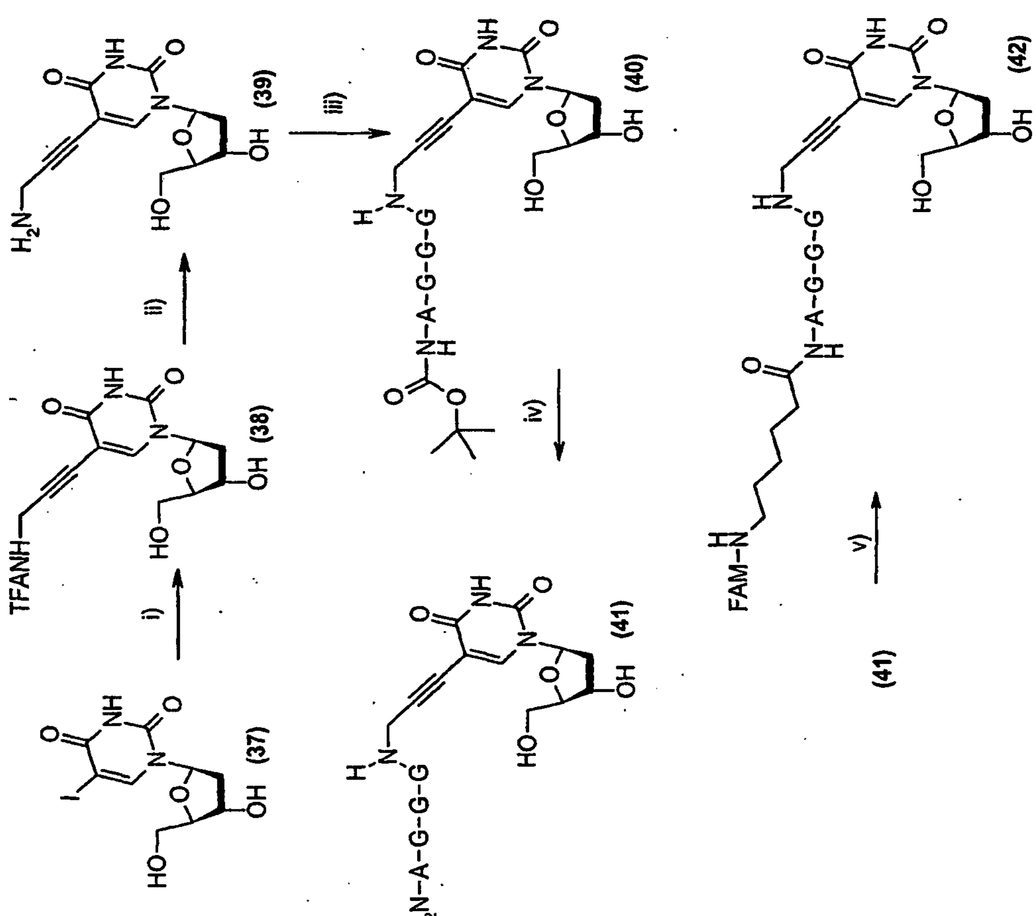


FIGURE 12

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 01/02402

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07H19/10 C07H19/20 C07H21/00 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07H C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, CHEM ABS Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PROBER J M ET AL: "A SYSTEM FOR RAPID DNA SEQUENCING WITH FLUORESCENT CHAIN-TERMINATING DIDEOXYNUCLEOTIDES" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, US, vol. 238, no. 4825, 16 October 1987 (1987-10-16), pages 336-341, XP000604017 ISSN: 0036-8075	1
A	the whole document	17

	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 September 2001	10/09/2001
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Authorized officer de Nooy, A

2

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	CONFALONE P N: "The use of heterocyclic chemistry in the synthesis of natural and unnatural products" JOURNAL OF HETEROCYCLIC CHEMISTRY, HETEROCORPORATION. PROVO, US, vol. 27, 1990, pages 31-46, XP002159718 ISSN: 0022-152X page 42 -page 45	1
X	CASALNUOVO A L ET AL: "PALLADIUM-CATALYZED ALKYLATIONS IN AQUEOUS MEDIA" JOURNAL OF THE AMERICAN CHEMICAL SOCIETY, AMERICAN CHEMICAL SOCIETY, WASHINGTON, DC, US, vol. 112, no. 11, 23 May 1990 (1990-05-23), pages 4324-4330, XP000561945 ISSN: 0002-7863 page 4330	1
A	SAUER M ET AL: "DETECTION AND IDENTIFICATION OF SINGLE DYE LABELED MONONUCLEOTIDE MOLECULES RELEASED FROM AN OPTICAL FIBER IN A MICROCAPILLARY: FIRSTSTEPS TOWARDS A NEW SINGLE MOLECULE DNA SEQUENCING TECHNIQUE" PHYSICAL CHEMISTRY CHEMICAL PHYSICS, ROYAL SOCIETY OF CHEMISTRY, CAMBRIDGE, GB, vol. 1, no. 10, 15 May 1999 (1999-05-15), pages 2471-2477, XP000854111 ISSN: 1463-9076 page 2474	2
A	US 5 681 940 A (WANG GUANGYI ET AL) 28 October 1997 (1997-10-28) cited in the application figures	10
A	US 5 268 486 A (MUJUMDAR RATNAKAR B ET AL) 7 December 1993 (1993-12-07) cited in the application claims	8

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 1 (in part)

Present claim 1 relate to an extremely large number of possible compounds. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to the compounds which fall within the structure of Formula I of claim 2.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Internat Application No
PCT/GB 01/02402

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5681940 A	28-10-1997	AU 690394 B	23-04-1998
		AU 4152596 A	31-05-1996
		CA 2202280 A	17-05-1996
		CN 1170412 A	14-01-1998
		CZ 9701291 A	15-10-1997
		EP 0789706 A	20-08-1997
		HK 1007881 A	30-04-1999
		HU 77516 A	28-05-1998
		JP 10506915 T	07-07-1998
		PL 319944 A	01-09-1997
		RU 2145964 C	27-02-2000
		SI 9520113 A	30-06-1998
		SK 54897 A	10-09-1997
		WO 9614329 A	17-05-1996
		US 5712378 A	27-01-1998
		US 6191266 B	20-02-2001
US 5268486 A	07-12-1993	US 5486616 A	23-01-1996
		US 5569766 A	29-10-1996
		DE 3912046 A	15-03-1990
		DE 3943817 C	12-07-2001
		JP 2191674 A	27-07-1990
		JP 2757965 B	25-05-1998
		JP 10096727 A	14-04-1998
		JP 2898264 B	31-05-1999
		JP 10088012 A	07-04-1998
		US 5569587 A	29-10-1996
		US 6225050 B	01-05-2001
		US 5627027 A	06-05-1997
		US 6048982 A	11-04-2000